



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, 15/82, 15/63, 5/10, C07K 14/435, A01N 63/02, A01H 1/00, A01K 67/027, C07K 1/12	A1	(11) International Publication Number: WO 95/29235 (43) International Publication Date: 2 November 1995 (02.11.95)
(21) International Application Number: PCT/GB95/00917 (22) International Filing Date: 24 April 1995 (24.04.95) (30) Priority Data: 9408466.2 27 April 1994 (27.04.94) GB (71) Applicant (for all designated States except US): BRITISH TECHNOLOGY GROUP LIMITED [GB/GB]; 101 Newington Causeway, London SE1 6BU (GB). (72) Inventors; and (75) Inventors/Applicants (for US only): BELL, David, Robert [GB/GB]; The University of Nottingham, University Park, Nottingham NG7 2RD (GB). USHERWOOD, Peter, Norman, Russell [GB/GB]; The University of Nottingham, University Park, Nottingham NG7 2RD (GB). DULUBOVA, Irina [RU/US]; The Rockefeller University, 1230 York Avenue, New York, NY 10021-6399 (US). VOLKOVA, Tatiana [RU/RU]; Russian Academy of Sciences, ul. Miklukho-Maklaya, Moscow, 117871 (RU). GRISHIN, Eugene [RU/RU]; Shemyakin & Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya, 16/10, Moscow, 117871 (RU). KRASNOPEROV, Valery [RU/RU]; Nelson		Institute of Environmental Medicine, London Meadow Lane, Tuxedo, NY 10987 (RU). GALKINA, Tatiana Genrikhovna [RU/RU]; Shemyakin & Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya, 16/10, Moscow, 117871 (RU). KHOVOTCHEV, Mikhail Vladimirovich [RU/RU]; Shemyakin & Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya, 16/10, Moscow, 117871 (RU). PLUZHNIKOV, Kirill Andreevich [RU/RU]; Shemyakin & Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya, 16/10, Moscow, 117871 (RU). SHAMOTENKO, Oleg Grigorievich [RU/GB]; Imperial College of Science Technology and Medicine, Wolfson Laboratories, Dept. of Chemistry, London CW7 2AY (GB). (74) Agent: SKINNER, Michael, Paul; Swindell & Pearson, 48 Friar Gate, Derby DE1 1GY (GB). (81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG). Published <i>With international search report.</i> <i>With amended claims.</i>
(54) Title: SPIDER NEUROTOXINS AND METHOD OF PRODUCING THE SAME		
(57) Abstract <p>A toxin comprising an isolated derivative or analogue of an invertebrate specific neurotoxin, δ-Latroinsectotoxin (δ-LIT). The toxin is formed by expressing, in a bacterial host, a nucleotide sequence corresponding to a truncated form of a gene from the genome of the Black Widow Spider. The gene encodes for a non-toxic precursor protein, whilst the truncated form encodes for an active toxin.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

Spider neurotoxins and method of producing the same

The present invention relates to a novel toxin, and a method of producing a toxin, particularly but not exclusively to an insect specific neurotoxin

δ -Latroinsectotoxin (δ -LIT), and a method of producing same.

A family of high molecular weight neurotoxins has been found in the venom of the black widow spider (*Latrodectus mactans Treddecimguttatus*). Some of these toxins have been identified as being either vertebrate or invertebrate specific. α -Latrotoxin (α -LT) and α -Latroinsectotoxins (α -LIT) are two such neurotoxins that have been characterised as being vertebrate and invertebrate specific respectively. The primary structures of these proteins have been determined, but characterisation of the structural features of the cloned toxins has not been possible due to the inability to achieve functional expression of their genes.

It is an object of the present invention to provide a novel toxin and a method of producing a toxin usually naturally produced by post-translational modification of a precursor protein, using recombinant technology.

According to the present invention there is provided a polypeptide, such as a toxin, formed by expression of a truncated form of a gene sequence, or an

- 2 -

analogue thereof.

Preferably the polypeptide is a neurotoxin and preferably corresponds to a toxic derivative of a substantially non-toxic precursor polypeptide encoded by the gene sequence. The polypeptide may comprise an amino acid sequence that corresponds to a truncated form of the amino acid sequence of a substantially non-toxic precursor polypeptide. Preferably the amino acid sequence of the polypeptide corresponds to the amino acid sequence of the precursor polypeptide with truncation thereof principally at the carboxy (C) end, and desirably by about 150 to 200 amino acids. The polypeptide amino acid sequence may in addition correspond to the precursor polypeptide amino acid sequence truncated at the amino end (N) preferably by less than 50 amino acids, and desirably by 7 or 28 amino acids.

The amino acid sequence of the polypeptide may be homologous to the amino acid sequence of the insect specific neurotoxin δ -Latroinsectotoxin (δ -LIT) or an active derivative thereof, and preferably comprises an amino acid sequence as shown in SEQIDN01 and SEQIDN02 or an active derivative thereof. Preferably the toxin is expressed from a nucleotide construct or truncated form of the gene sequence comprising a sequence as shown in SEQIDN01, or active variants thereof. Preferably the

- 3 -

toxin is expressed from a sequence substantially as provided in a microorganism deposited at The National Collections of Industrial and Marine Bacteria Limited, under Accession No. NCIMB 40632.

The invention also provides a protein for use as a toxin comprising an amino acid sequence substantially as shown in SEQIDN01 and SEQIDN02, or an active derivative thereof.

According to a further aspect of the present invention there is provided a nucleotide sequence comprising a truncated form of a gene sequence or an analogue thereof, for use in the expression of a polypeptide, such as a toxin.

Preferably the nucleotide sequence corresponds to a gene encoding for a precursor polypeptide and truncated at the 3' end thereof or an active derivative thereof. Preferably the nucleotide sequence corresponds to the gene truncated by about 400 to 650 nucleotide bases, and desirably between 550 to 600 nucleotide bases.

The nucleotide sequence may also correspond to the gene truncated at the 5' thereof, preferably by less than 100 nucleotide bases, and desirably by either 84 or 21 nucleotide bases.

- 4 -

Preferably the nucleotide sequence corresponds to part of a gene encoding for a neurotoxin in the venom of the Black Widow Spider (*Latrodectus mactans* *Tredecimguttatus*), or an active derivative thereof.

The nucleotide sequence may correspond to part of the gene encoding the precursor polypeptide of insect specific toxin δ -Lactoinsectotoxin (δ -LIT), or an active derivative thereof. The nucleotide sequence preferably codes for a polypeptide comprising a sequence of 991 amino acids.

Preferably the nucleotide sequence comprises a base sequence as shown in SEQIDN01, or an active derivative thereof, and preferably as comprised in a microorganism deposited under Accession No. NCIMB 40632 at The National Collections of Industrial and Marine Bacteria Limited.

Preferably the nucleotide sequence codes for a polypeptide having an amino acid sequence as shown in SEQIDN01 and SEQIDN02, or an active derivative thereof.

The nucleotide sequence may be a cDNA derived from mRNA by the use of an enzyme such as reverse transcriptase. The nucleotide sequence may alternatively be an oligonucleotide DNA construct produced perhaps using the polymerase chain reaction (PCR).

- 5 -

According to a further aspect of the present invention there is provided a method of producing a polypeptide, the method comprising producing a recombinant DNA molecule comprising a truncated form of a gene, and expressing the truncated form in a host expression system, such as a viral or bacterial expression system, to produce the polypeptide.

Preferably the polypeptide produced is an active toxin and desirably a neurotoxin substantially as defined above. Preferably the truncated form comprises part of a gene which encodes for a non-toxic precursor polypeptide.

Preferably the truncated form comprises a nucleotide sequence substantially as defined above, and as shown in SEQIDNO1, or an active derivative thereof. Preferably the expression system comprises E.coli BL21 (DE3) bacterial cells transformed with pT7-7 vectors comprising the truncated form of the sequence, desirably substantially as deposited under Accession No. NCIMB 40632 at The National Collections of Industrial and Marine Bacteria Limited. The expression system may comprise a baculovirus system.

In a further aspect of the present invention there is provided a recombinant DNA molecule, such as a virus, and in particular a baculovirus comprising a truncated form of a gene encoding for a toxin generally

- 6 -

as defined above, and substantially as provided in the microorganism deposited under Accession No. NCIMB 40632.

A still further aspect of the present invention provides an expression vector comprising a truncated form of a gene, the truncated form encoding for a toxin generally as defined above.

The invention also provides a cell, such as a viral or bacterial cell transformed with a recombinant molecule as defined above.

There is also provided an insecticide comprising a toxin as defined above. The insecticide may be so as to be administered orally or topically. The insecticide may comprise a spray.

This invention also provides an insecticide system comprising means for expressing a truncated form of a gene to produce a toxin as described above in an insect to kill or incapacitate the insect. The insecticide system may comprise a viral expression system, and desirably a baculovirus expression system.

According to a further aspect there is provided a plant comprising a genetically modified cell containing a truncated form of a gene sequence substantially as defined above.

Still further according to the present invention

- 7 -

there is provided a non-human animal comprising a genetically modified cell containing a truncated form of a gene sequence substantially as defined above.

According to a further aspect of the present invention there is provided a toxin formed by processing of a substantially isolated non-toxic precursor polypeptide.

The toxin is preferably a neurotoxin and is preferably formed by truncation toward the carboxy (C) end of the precursor polypeptide, preferably by site-directed mutagenesis. Desirably the toxin amino acid sequence generally corresponds to the amino acid sequence of the precursor polypeptide, truncated by between 150 and 200 amino acids. The toxin amino acid sequence may also be formed by truncation toward the amino (N) end of the precursor polypeptide amino acid sequence, the fragment cleaved therefrom preferably being significantly smaller than the fragment cleaved from the carboxy end, and may comprise 7 or 28 amino acids.

Preferably the toxin has an amino acid sequence corresponding to polypeptide encoded by part of a gene of the Black Widow Spider (*Latrodectus mactans* *Tredecimguttatus*). The toxin may comprise or be an analogue of the insect specific neurotoxin δ -Latroinsectotoxin (δ -LIT), or an active derivative

- 8 -

thereof.

Preferably the toxin comprises an amino acid sequence as shown in SEQIDN01 and SEQIDN02 or an active derivative thereof.

In a further aspect of the present invention there is provided a method of producing an active polypeptide from an inactive precursor polypeptide, the method comprising truncating the isolated precursor polypeptide.

Preferably the isolated precursor polypeptide is truncated at the Carboxyl end, perhaps using proteolytic cleavage, and preferably by site directed mutagenesis. Truncation of the N terminus may also be provided. Preferably the active polypeptide is a toxin and is substantially as described above.

According to another aspect of the present invention there is provided an isolated nucleotide base sequence encoding for a toxin precursor polypeptide as defined above and preferably with an amino acid sequence as shown in SEQIDN04 or an active derivative thereof. The base sequence preferably comprises the sequence shown in SEQIDN03 or a derivative thereof. The nucleotide base sequence preferably encodes a precursor polypeptide of the neurotoxin δ -Latroinsectotoxin (δ -LIT). Preferably the base sequence is substantially

- 9 -

as provided in the microorganism deposited under
Accession No. NCIMB 40633.

In a further aspect there is provided a recombinant
DNA molecule such as a virus, and more particularly a
baculovirus comprising a sequence as defined in the
preceding paragraph.

In a still further aspect the invention provides a
cell, such as a bacterial cell or viral cell, transformed
with a recombinant DNA molecule as described in the
preceding paragraph.

This invention also provides an insecticide system
comprising means for expressing a gene as described above
to produce a precursor polypeptide as described above and
to process the precursor polypeptide to produce a toxin
in an insect to kill or incapacitate the insect. The
insecticide system may comprise a viral expression
system, and desirably a baculovirus expression system.

According to a further aspect there is provided a
plant comprising a genetically modified cell containing a
gene as defined above.

Still further according to the present invention
there is provided a non-human animal comprising a gene-
tically modified cell containing a gene as defined above.

- 10 -

Preferred embodiments of the present invention will now be described by way of example only, with reference to the accompanying sequences, in which:-

SEQ ID NO. 1 shows the nucleotide base sequence and the corresponding amino acid sequence of a truncated form of a gene and a polypeptide encoded thereby, according to one aspect of the present invention;

SEQ ID NO. 2 shows the polypeptide sequence of SEQIDN01;

SEQ ID NO. 3 shows the nucleotide base sequence and the corresponding amino acid sequence of a gene and a polypeptide encoded thereby, according to another aspect of the present invention; and

SEQ ID NO. 4 shows the polypeptide sequence of SEQIDN03.

Referring to the sequences, a polypeptide such as a toxin as in SEQIDN02 is formed by expression of a truncated form of a gene sequence (SEQIDN01), or an analogue thereof.

A toxin from Black Widow Spider (*Latrodectus mactans Tredecimguttatus*) venom (BWSV),

- 11 -

δ -Latroinsectotoxin, (δ -LIT) has been purified and shown to possess insect specific toxicity. The δ -LIT structural gene has been cloned and sequenced and the N- and C termini of the native (precursor) and functional protein toxin have been determined as described below. Site directed mutagenesis of δ -LIT cDNA enabled expression of the mature protein product (toxin) in bacteria, and this has been shown to be toxic to locusts.

Expression and production of this and other such toxins in bacterial expression systems has hitherto not been possible. The invention includes identification of the sites for cleavage of the precursor protein to produce the toxin, and the precise site of truncation of the gene sequence which has enabled the toxin to be expressed in bacterial, and indeed other suitable hosts.

Microorganism deposits have been made under the Budapest Treaty on 3rd May 1994, at the National Collections of Industrial and Marine Bacteria Limited, of 23 St. Machar Drive, Aberdeen, Scotland, United Kingdom. Escherichia coli (XL-1 Blue pT7. δ M) cloned with the truncated form of the gene sequence is deposited under Accession No. 40632, and Escherichia coli (HMS 174 pT7. δ FL) cloned with substantially the full gene sequence is deposited under Accession No. 40633).

In more detail, the cDNA cloning and sequencing was

- 12 -

conducted as follows. Poly(A⁺)-RNA was isolated from venom glands of the Black Widow Spider (*Latrodectus mactans* *Tredecimguttatus*) and a cDNA library constructed in the plasmid vector pSP65 (according to Kiyatkin et al, 1993). A library of 6×10^4 clones was screened with an end-labelled 23-mer oligonucleotide probe based on the N-terminal sequence of δ -LIT (amino acid residues 1-8)-
5' GA(C/T)GA(A/G)GA(A/G)GA(C/T)GG(A/T)GAAATGAC 3' .
Hybridization was performed. Positive clones were colony-purified and analysed by restriction mapping. The inserts were excised and fragmented by sonication as described (Sambrook et al, 1989) followed by cloning into the SmaI site of pBluescript II SK+ and SK- vectors (Stratagene, USA). Single-stranded templates for sequencing were obtained after infection with helper phage VCS (Stratagene). The DNA sequences were determined by the chain-termination method (Sanger et al, 1977) using Sequenase 2.0 version kit (USB Corporation) and T7 and T3 vector-specific primers (Stratagene). Each sequence was determined at least twice on both strands. Synthetic primers were used to sequence regions that were not covered by isolated subcloned fragments.

DNA and protein sequence analysis was performed using the computer software DNASTAR (Dnastar Inc) and PCGENE (IntelliGenetics Inc). This work benefitted from the GCG programme mounted on the SERC Daresbury SEQNET

- 13 -

facility (Devereux, Haeberli and Smithies, (1984),
Nucleic Acids Research 12(1); 387-395).

The full-length cDNA construction was carried out as follows. Two sets of oligonucleotide primers were used to produce N- and C- overlapping parts of δ -LIT coding sequences by polymerase chain reaction. To facilitate subcloning into the expression vector the 5' sense primer (P1) (TTGGGATCCGATGAAGAAGATGGAGAA) and 3' antisense primer (P8) (CAATGGTCGACACAGAAGGAATGGTA) contained BamHI and SalI restriction enzyme sites. Two other primers -P9, sense (GTCTGAACCATTTACTGTCC) (position 1283-1302) and P3, antisense (GTAAGATTACCATCTGCAAC) (complementary to position 2253-2272) were chosen to produce overlapping fragments with an internal NcoI (2056) restriction site. An oligonucleotide was designed to terminate the protein sequence after amino acid 991-5' CGTTTCGTCGACTCATTCGGTAAAGTACGACGAAA 3' . The polymerase chain reaction was performed using 1 unit of Taq-polymerase (Promega) under standard conditions (30 cycles, 55°C for 1 min, 72°C for 3 min, 94°C for 1 min, with 100 pmol of each primer and 1-10 ng first-strand cDNA). In the first cycle the denaturation time was elongated to 5 min. The molecular mass of the amplified material was checked on an agarose gel. First-strand cDNA synthesis was carried out using First-strand cDNA Synthesis Kit (Pharmacia) with both random and specific

- 14 -

primers as recommended by the manufacturer. The PCR products were purified from agarose gel using GeneClean Kit (Bio 101 Inc.), digested with appropriate pairs of enzymes (BamHI and NdeI for the N-terminus part and SalI/NdeI for the C-terminus) and cloned into the pT7-7 vector restricted with the similar pairs of enzymes. The full-length cDNA was created as a result of three-way ligation between N-terminal BamHI/NdeI-fragment, C-terminal NdeI/SalI-fragment and pT7-7 BamHI/SalI-digested vector. The final construct had eight additional amino acid residues at the amino terminal end (MARIRARG). All plasmid constructs were verified by sequencing from both ends and through the junction region. The full length construct was designated pT7 δ .FL a sample of which is deposited at the NCIMB, accession No 40633 and the truncated clone (1-991 amino acids) was designated pT7. δ M. (NCIMB No 40632).

In order to verify the identity of the δ -LIT cDNA, this clone was expressed in the bacterial pT7-7 vector in E.coli BL21(DE3) cells. A full-length toxin cDNA (corresponding to Asp residue 29 to 1186) 1214 of SEQION04 was constructed and designated pT7. δ FL. The first 28 amino acids are believed to be present in the precursor polypeptide in spider venom glands, but cleaved during N-terminal processing. The recombinant protein constitutes approximately 10% of the total bacterial

lysate protein. A polyclonal antibody specific was raised to δ -LIT purified from spider venom glands, and demonstrated to be specific for the δ toxin. This protein specifically detected a protein of 130 kDa in bacteria expressing recombinant full-length δ -LIT. Comparison of the molecular mass of the bacterially expressed full-length δ -LIT and the toxin purified from venom glands demonstrated a size difference of approximately 23kDa, in agreement with the calculated molecular mass. The full-length δ -LIT had no toxicity towards insects and is considered to be an inactive precursor form of the toxin.

δ -LIT purified from venom glands was analysed by mass spectrometry (on a Kratos Kompact MALDI 3 Mass Spectrometer, using sinapinic acid as a matrix. The nitrogen laser excitation was at 337nm, and the positive ion was detected in the linear mode) yielding a prominent molecular ion with a $m/z+$ ratio of 110916. This corresponds closely to the expected molecular mass of δ -LIT which is truncated at amino acid 991. By comparison, the bacterially expressed full length δ -LIT yielded a molecular ion with an $m/z+$ ratio of 133631 (VK, DRB, PNRU, Data not shown), within 100 Da of the calculated value. Site directed mutagenesis was used to create a novel δ -LIT cDNA clone (pT7 δ M), which was truncated after amino acid 991 of the δ -LIT sequence

- 16 -

(SEQIDN02). This protein was expressed in bacteria, yielding a protein of similar molecular mass to the mature toxin isolated from spider venom.

E. coli BL21(DE3) cells transformed with pT7 clones were grown in LB medium containing 100mg ampicillin/ml at 30°C to an A_{600nm} of approximately 0.5. Then expression was induced by addition of IPTG (1mM) to the medium, and incubation continued for 1 hour. For functional studies, bacteria were washed and resuspended in 50 mM TrisHCl, 100mM NaCl, 10mM KCl, 0.4% Triton X-100, 12% (W/V) sucrose, 5mM DTT, 2 μ g/ml aprotinin, 2mM EDTA, pH8, and sonicated on ice. Ammonium sulphate was added to the cleared supernatant to a final concentration of 20% of saturation, and the pellet was resuspended in buffer without DTT. These samples (5-15 μ l) were used for thoracic injection into locusts (100-300 mg body weight); each test was performed on more than 4 locusts, and the locusts were examined for toxicity for 24 hours. Extracts from pT7-7 and pT7.6FL produced no effects on the locusts, but extracts from bacteria carrying pT7.6M caused rapid lethality. The time of death of the locusts varied from 5 minutes - 4 hours, depending on the potency of the batch of toxin.

Preliminary studies were undertaken on

neurally-excited and resting retractor unguis nerve-muscle preparations isolated from metathoracic legs of adult (male and female) locusts (Usherwood and Machili, 1968). δ -LIT was applied in standard locust saline (mM: NaCl, 180; KCl, 10; CaCl₂, 2; Hepes, 10 (pH 6.8)). A few studies were undertaken using saline in which CaCl₂ was omitted. Mechanical responses were recorded using a Grass strain gauge connected to a Grass pen recorder. Recordings of miniature excitatory postsynaptic potentials were made from fibres of metathoracic extensor tibiae muscles of adult locusts (either sex) using intracellular microelectrodes (approximately 10M Ω resistance). δ -LIT was applied in either standard locust saline, saline in which CaCl₂ was omitted or saline which contained MgCl₂ substituted for NaCl. The miniature potentials were recorded on video tape and analysed on a MassComp computer using in-house software. Membrane bilayers were formed at the tips of patch pipettes (diam 1-2 μ m; fabricated from Clark Electromedical glass) from monolayers of either diphytanoyl phosphatidylcholine or a mixture of 9 parts isolectin and 1 part cholesterol using a pipette dipping technique (Montal and Muller, 198). Similar patch pipettes were used to excise membrane patches from locust metathoracic extensor tibiae muscle fibres (Huddle et al). In order to reduce the activities of endogenous potassium channels KCl was eliminated from the pipette

- 18 -

and bath salines.

The neurally-evoked twitch contraction of the locust retractor unguis muscle was reduced by approximately 40% by 10^{-11} M \mathcal{S} -LIT (applied in standard saline) and was abolished during application of 10^{-10} M toxin. Small spontaneous contractions sometimes occurred during \mathcal{S} -LIT application. The changes in twitch amplitude were accompanied by an irreversible muscle contracture. The appearance of the contracture was delayed and its amplitude was reduced when the concentration of \mathcal{S} -LIT was lowered. A muscle contracture also occurred when toxin was applied when the muscle was not neurally stimulated. Twitch contractions do not occur in calcium-free saline and when 10^{-10} M toxin was applied to a preparation equilibrated in this saline a contracture did not occur even after 30 min application of the toxin.

When inside-out patches excised from locust muscle fibres were exposed to 10^{-11} M α -LIT in the patch pipette, channel opening, of maximum conductance approximately 40 pS, were observed. Channel openings of this type were never seen in the absence of toxin. The channel current exhibited inward rectification when the patch pipette and bath contained identical salines (including 2 mM CaCl_2), and channel open times were longer at negative than at positive pipette potentials. When

- 19 -

there was a 10-fold Ca^{2+} gradient across a patch, the reversal potential of the channel current was $\pm 15\text{mV}$, the sign being dependent on the Ca^{2+} gradient.

In the artificial bilayer studies where 10^{-11}M ω -LIT was placed in the patch pipettes, single channel openings of approximately 30pS conductance were observed. These channels were not seen when toxin was omitted from the patch pipette. With identical salines (containing 2mM CaCl_2) in the patch pipette and bath, the current-voltage characteristic of the ω -LITx channel was sigmoidal with a reversal potential at 0mV. The channel was shown to be Ca-selective by manipulating the ionic regimes of patch pipette and bath.

A cDNA library from venom gland cDNA was screened with a 23-bp oligonucleotide probe corresponding to the N-terminal sequence of ω -LIT (as described above). To reduce the number of nucleotide ambiguities the codon usage data available from the nucleotide sequences of ω -LT and ω -LIT cDNA (Kiyatkin et al, 1990, Kiyatkin et al 1993) was referred to. Five positive cDNA clones were colony-purified and sequenced. The longest clone (pDT-1) contained more than 2 (kb) of ω -LIT coding region. A PstI-3' fragment was used to rescreen the cDNA library to search for clones encoding the C-terminal part of the toxin. An additional cDNA clone, pDT-17, was isolated,

- 20 -

which covered the C-terminal coding region of the δ -LIT. Two overlapping clones, covering the entire open reading frame, have been sequenced in their entirety. The two clones have been demonstrated to be part of a single, continuous RNA from venom glands by polymerase chain reaction across the overlapping region, using two distinct sets of primers. The composite clones encode a cDNA with a frame of 3642 bp starting from the first in-frame Methionine and ending with TAA stop codon (SEQIDN03).

The Met residue is preceded by an in-frame stop codon confirming the full length of the deduced sequence.

δ -LIT was purified to homogeneity from Black Widow Spider venom by three rounds of column chromatography according to (Krasnoperov et al, 1992). 23 amino acid residues of the N-terminal sequence of δ -LIT was sequenced. The pure toxin was digested with trypsin and seven individual peptides were isolated and partially sequenced.

Direct N-terminal sequence determination demonstrates that the mature protein starts from the sequence DEEDGEM..., so residue 1 in SEQIDN01 and 2 is the first Asp of this sequence. The deduced polypeptide starting from Asp (+1) consists of 1186 amino acid (as

- 21 -

shown in SEQIDNO3⁺⁴, Asp residue 29 to residue 1214) residues with a predicted molecular mass of 132671 Daltons and pI of 5.4. It contains all of the peptide sequences determined by amino acid sequencing analysis. There are two in-frame Met residues (-7 and -28) upstream of the N-terminus (as shown in SEQIDNO3) of the mature protein which can serve as translation initiation sites. The nucleotide sequence surrounding the ATG codon for Met (-7) correlates better with the classical Kozak consensus (Kozak, 1989), but the nucleotide arrangement for Met (-28) strongly corresponds to starting points for at least two other known proteins which have been isolated from arachnids: Major house dust mite allergen (AAAATGA) (Yuuki et al, 1991) and Low molecular weight protein co-purified with α -Latrotoxin (AAATGA) (Kiyatkin et al, 1992). In both cases, the deduced sequence preceding the N-terminus of the mature protein does not correspond to classical signal peptide structures. We conclude that post-translational modification of δ -LIT N-terminus is limited to removal of 7 or 28 amino acid residues. The existence of a cluster of positive amino acid residues Arg-X-Lys-Arg (-1-4) which can serve as a potential endopeptidase-cleavage site supports the hypothesis that post-translational processing occurs at the N-terminus.

Analysis of the deduced structure of δ -LIT with PEST (Rogers, S. et al, 1986) reveals the presence of an

- 22 -

amino acid sequence enriched in P, E, S and T, which has previously been correlated with rapid degradation of intracellular proteins (Gottesman & Maurizi, 1992). This region has the sequence EESGAPEGSF DSPSS, and is situated between residues 956-970. The presence of the PEST-region in the C-terminal part of δ -LIT is consistent with C-terminal processing of this protein.

Computer analysis of δ -LIT predicts three putative transmembrane helixes two of them situating in terminal regions (residues 39-67 and 221-240) and the third one of a minimal length (residues 580-595) being in the central region. The second putative transmembrane helix (residues 221-240) belongs to a very conservative region between all spider high molecular weight protein neurotoxins (Kiyatkin et al, 1993).

Dot-matrix analysis of the predicted δ -LIT amino acid sequence revealed the presence of a repeated motif in the central part of the protein molecule. 460 amino acid residues of the δ -LIT primary structure comprise tandemly arranged imperfect copies of the ankyrin-like repeats (Michaely & Bennett, 1992). Whereas α -LT and α -LIT (Kiyatkin et al, 1990, Kiyatkin et al, 1993) have no less than 20 repeats, δ -LIT has been found to have only 13 successive repeated units. Their optimal alignment is with phasing originally suggested in (Lux et

al, 1990). The sequence of 13 amino acids which precede the first repeat can be viewed as a reduced repeated unit according to its good correlation with a consensus sequence. The majority of δ -LIT repeated units are 33-34 amino acids in length, but two repeats contain 35 (R1) and 36 (R6) residues, respectively.

Analysis with the PCOMPARE programme showed the linear correlation between the repeated units of two insect-specific toxins. Strong linear correspondence has been found for δ -LIT repeats R2-R9 in comparison to the analogous repeats in α -LIT (Kiyatkin et al, 1993). The first repeat in δ -LIT does not correspond well to the first one in α -LIT and shows high similarity to R7 from α -LIT. δ -LIT repeat R10 is most similar to R19 from α -LIT: this repeat is unusual in having Ser and Gly residues at position 8 and 31, respectively. The next stretch of similarity is found between R11-R13 of δ -LIT and R10-R12 of α -LIT. We have noted that the R7, R2 and R9 repeats are the most highly conserved between the insectotoxins, suggesting a functional role in insectotoxicity. It has been shown that Erythrocyte Ankyrin repeats are not equivalent in respect of their functional ability to bind different proteins (Davis et al, 1991), and thus toxin repeats are also expected to make different contributions to their function.

- 24 -

Dot-matrix comparison of δ - and α -LIT shows that they share a similar overall organization, with the strong central diagonal broken once (between 900 and 1130 amino acid residues of δ -LIT) and restored for the last 160 amino acids of both toxins. The displacement of the central diagonal reflects the difference in toxin length; δ -LIT is 190 amino acids shorter than its insect-specific counterpart.

The mature protein can be divided into several structural domains: an N-terminus consisting of about 470 amino acid residues and possessing strong linear homology with α -LIT; the central domain of about 430 amino acids almost completely comprised of tandemly arranged ankyrin-like repeated units and a C-terminal domain of about 160 amino acids.

Alignment of the insectotoxin protein sequences shows that both the N- and C-terminal structural domains demonstrate the presence of high identity regions separated by rather dissimilar sequences, with a high level of identity (44.9% for the N-terminal domain and 37.1% for the C-terminus). The most dramatic changes in primary structure of the two insectotoxins are concentrated in C-terminal parts of the repeat containing domains. A stretch of homology is localized to 13 ankyrin repeat units of δ -LIT. This region is followed

- 25 -

by a sequence of about 110 amino acid residues that has no obvious homology either with α -LIT or α LI nor with any other proteins from NBRF-PIR database.

Interestingly, this domain, which is absent from δ -LIT, forms a specific region in the primary structure of α -LIT that has an unusual clustering of Cys-residues and possesses homology with mammalian-specific α LI (Kiyatkin et al, 1993). So striking structural difference between the two insect-specific neurotoxins suggests that the C-terminal part of the ankyrin-like repeated domain plays a particular role in providing a structural basis for their different functional properties.

The high molecular weight protein toxins from the venom of the Black Widow Spider are a potent and specific class of toxins. These toxins offer a great potential for elucidating the function of neural proteins, and for providing insect specific toxins. However this potential has not previously been realised due to the inability to express these protein toxins with any function. The present invention provides for the cloning of a novel DNA transcript encoding for a novel insect-specific toxin, and functional expression of this toxin, and other polypeptides in bacteria.

The δ -LIT cDNA was cloned with an oligonucleotide based on the sequence of amino acids 1-23 of the toxin,

- 26 -

and its identity confirmed by additional peptide sequences, and immunochemical identity, using an antibody specific for the \mathcal{S} -LIT. The deduced primary structure of \mathcal{S} -LIT has considerable similarity to the mammalian specific α LT and the insect-specific α -LIT, suggesting that these toxins are part of a family with similar structure. The three proteins have a central domain which is composed of "ankyrin-like" repeats, with 13 repeats in \mathcal{S} -LIT. The ankyrin family of proteins couple spectrin to a variety of integral membrane proteins (Bennett, 1992), and it is believed that the "ankyrin repeat" domain of the ankyrins is responsible for specific binding to proteins (Davis and Bennett, 1990 J Biol Chem 265: 10589-10596; Davis et al (1991) J. Biol Chem 266: 11163-11169). This structural similarity with the ankyrin family is reflected in the known functional properties of the latrotoxins; α LT is known to bind to a receptor with high affinity (K_d 10^{-9} M). It remains to be determined whether this specific binding to the α LT receptor is mediated via the ankyrin repeat region of the toxin.

Surprisingly, \mathcal{S} -LIT has no greater similarity to the insect-specific α -LIT (38%) than to the mammalian-specific α LT (37%). Whereas \mathcal{S} -LIT has only 13 repeats, the α LT and α -LIT have 19 and 20 ankyrin repeats, respectively. The latter 6/7 repeats have no counterpart

- 27 -

in the δ -LIT, and may be a structural unit, as the α -toxins both contain 6 cysteine residues in this region, with partially conserved spacing. However, in view of the differences in target toxicity of the α LI and α -LIT, it is not possible to identify this structural features with insect-specific toxicity.

δ -LIT exhibits a marked disparity between the molecular weight of the toxin, as deduced from the cDNA sequence, and the relative mobility of the pure toxin purified from venom. Whilst the N-terminus of the protein has been identified unambiguously by protein sequencing, the precise position of the C-terminus has been difficult to document. Expression of the full length δ -LIT cDNA in bacteria demonstrated that the calculated molecular mass is accurately reflected in the relative mobility of the protein on SDS-PAGE, and that the natural venom derives predominantly, if not wholly from proteolytic, C-terminal processing. The full-length recombinant protein was purified, but was not toxic to locusts under any conditions. The full-length protein is an inactive precursor of the functional toxin.

The precise site of the C-terminus of δ -LIT purified from venom was assessed by MALDI-mass spectrometry, which localised the site of cleavage to amino acid 991 of the protein. The cDNA was mutated to

- 28 -

produce a protein of 991 amino acids with a sequence as shown in SEQIDNO2, and expressed in bacteria. The mature recombinant protein was soluble and was lethal to locusts. Partial purification of the protein suggests that the toxin is highly toxic.

Expression of the mature toxin from using the truncated form of the full gene sequence as described above has considerable advantages. Firstly, the toxin can be produced relatively easily by functional expression of the truncated form in a bacterial system, thereby obviating the need to purify toxin from venom glands of spiders. This enables industrial production of the toxin and hence commercial exploitation, for example as the major component of an insecticide system.

Moreover, it presents possible administration systems for the toxin as an insecticide, beside conventional methods such as spraying. For example it may be possible to produce a modified plant cell or plant, such as a crop plant, containing a recombinant molecule incorporating the truncated sequence. Such a system comprises a recombinant baculovirus comprising the truncated form. Such viruses are highly infectious in vivo and resistant to inactivation in host cells, and are capable of high levels of expression of the inserted nucleotide sequence in host insect cells. This is expected to be harmless to the plant and indeed to

- 29 -

vertebrates. Upon ingestion of the plant tissue an insect will take in the recombinant molecule and/or toxin, resulting ultimately in the death of the insect. Since the toxin is insect specific, it is expected to have no detrimental effect to humans or animals upon consumption.

This is an example of one use of one embodiment of the invention to express a toxin that is usually produced by post-translational modification of a precursor protein in biological systems, in a bacterial expression system. It is to be appreciated that the truncated form of other genes coding for other proteins could be expressed in this way, and fall within the scope of the present invention.

The invention also provides toxin formed from the expression of a full, isolated gene to produce a precursor polypeptide which is then post-translationally modified. The precursor polypeptide has an amino acid sequence as shown in SEQIDN04, and the toxin has a sequence as shown in SEQIDN02.

The isolated gene (SEQIDN03) (or an analogue) encoding for the precursor polypeptide of the toxin δ LIT can be cloned into a vector for expression of the precursor polypeptide. A baculovirus expression system

- 30 -

can be used. The precursor polypeptide thus produced can then be truncated at the sites indicated above, by site directed mutagenesis, to produce an active toxin. This enables the toxin δ LII or an active derivative thereof, to be produced independently of the Black Widow Spider, and thus on an industrial scale, for use as indicated as an insecticide.

Whilst endeavouring in the foregoing specification to draw attention to those features of the invention believed to be of particular importance it should be understood that the Applicant claims protection in respect of any patentable feature or combination of features hereinbefore referred to and/or shown in the drawings whether or not particular emphasis has been placed thereon.

- 31 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: BRITISH TECHNOLOGY GROUP LIMITED
- (B) STREET: 101 NEWINGTON CAUSEWAY
- (C) CITY: LONDON
- (E) COUNTRY: UNITED KINGDOM
- (F) POSTAL CODE (ZIP): SE1 6BU

(ii) TITLE OF INVENTION: A NOVEL TOXIN AND A METHOD OF PRODUCING A TOXIN

(iii) NUMBER OF SEQUENCES: 4

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release £1.0, Version £1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2976 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PLASMID DNA"

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: LATRODECTUS MACTANS TREDECIMGUTTATUS

(vii) IMMEDIATE SOURCE:

- (B) CLONE: pT7.deltam

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..2976

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GAT	GAA	GAA	GAT	GGA	GAA	ATG	ACT	CTA	GAA	GAA	AGA	CAA	GCA	CAA	TGC	48
Asp	Glu	Glu	Asp	Gly	Glu	Met	Thr	Leu	Glu	Glu	Arg	Gln	Ala	Gln	Cys	
1				5					10					15		
AAA	GCA	ATA	GAG	TAC	AGC	AAT	TCA	GTT	TTT	GGG	ATG	ATC	GCT	GAT	GTA	96
Lys	Ala	Ile	Glu	Tyr	Ser	Asn	Ser	Val	Phe	Gly	Met	Ile	Ala	Asp	Val	
			20					25					30			
GCT	AAC	GAC	ATC	GGT	TCC	ATT	CCT	GTA	ATT	GGC	GAA	GTA	GTT	GGC	ATT	144
Ala	Asn	Asp	Ile	Gly	Ser	Ile	Pro	Val	Ile	Gly	Glu	Val	Val	Gly	Ile	
			35				40					45				
GTA	ACT	GCC	CCA	ATT	GCC	ATC	GTA	AGT	CAC	ATT	ACT	AGC	GCA	GGC	TTG	192
Val	Thr	Ala	Pro	Ile	Ala	Ile	Val	Ser	His	Ile	Thr	Ser	Ala	Gly	Leu	
			50			55					60					
GAT	ATA	GCT	TCT	ACG	GCA	TTA	GAT	TGT	GAT	GAT	ATA	CCT	TTT	GAT	GAG	240
Asp	Ile	Ala	Ser	Thr	Ala	Leu	Asp	Cys	Asp	Asp	Ile	Pro	Phe	Asp	Glu	

65	70										- 32 -		75		80				
ATT Ile	AAG Lys	GAA Glu	ATA Ile	TTA Leu 85	GAA Glu	GAA Glu	AGA Arg	TTC Phe	AAT Asn 90	GAA Glu	ATA Ile	GAT Asp	AGA Arg	AAG Lys 95	TTG Leu	228			
GAC Asp	AAG Lys	AAC Asn	ACA Thr 100	GCT Ala	GCT Ala	TTG Leu	GAA Glu	GAG Glu 105	GTC Val	TCT Ser	AAA Lys	CTG Leu 110	GTA Val	AGT Ser	AAA Lys	336			
ACT Thr	TTT Phe	GTT Val 115	ACG Thr	GTG Val	GAA Glu	AAA Lys	ACA Thr 120	AGG Arg	AAT Asn	GAA Glu	ATG Met	AAC Asn 125	GAA Glu	AAT Asn	TTT Phe	384			
AAG Lys	CTT Leu 130	GTT Val	TTG Leu	GAA Glu	ACT Thr	ATA Ile 135	GAA Glu	AGC Ser	AAA Lys	GAA Glu	ATA Ile 140	AAA Lys	TCA Ser	ATT Ile	GTA Val	432			
TTT Phe 145	AAA Lys	ATA Ile	AAT Asn	GAT Asp	TTT Phe 150	AAA Lys	AAG Lys	TTT Phe	TTT Phe	GAA Glu 155	AAA Lys	GAA Glu	CGA Arg	CAA Gln	AGA Arg 160	480			
ATT Ile	AAA Lys	GGT Gly	TTG Leu	CCT Pro 165	AAA Lys	GAT Asp	AGG Arg	TAT Tyr	GTT Val 170	GCT Ala	AAG Lys	CTT Leu	CTA Leu	GAA Glu 175	CAA Gln	528			
AAA Lys	GGT Gly	ATT Ile 180	TTA Leu	GGT Gly	TCT Ser	TTA Leu	AAA Lys 185	GAA Glu	GTA Val	AGA Arg	GAA Glu	CCA Pro	TCT Ser 190	GGA Gly	AAC Asn	576			
AGT Ser	CTG Leu	AGC Ser 195	TCC Ser	GCG Ala	TTA Leu	AAT Asn	GAA Glu 200	CTC Leu	TTA Leu	GAC Asp	AAA Lys	AAC Asn 205	AAC Asn	AAC Asn	TAT Tyr	624			
GCC Ala 210	ATC Ile	CCA Pro	AAA Lys	GTG Val	GTT Val	GAT Asp 215	GAT Asp	AAT Asn	AAG Lys	GCC Ala	TTT Phe 220	CAG Gln	GCG Ala	CTG Leu	TAT Tyr	672			
GCT Ala 225	TTA Leu	TTT Phe	TAT Tyr	GGA Gly	ACT Thr 230	CAG Gln	ACT Thr	TAT Tyr	GCA Ala	GCC Ala 235	GTT Val	ATG Met	TTT Phe	TTT Phe	TTA Leu 240	720			
CTC Leu	GAA Glu	CAA Gln	CAT His 245	TCT Ser	TAT Tyr	CTG Leu	GCT Ala	GAT Asp	TAT Tyr 250	TAT Tyr	TAC Tyr	CAA Gln	AAA Lys	GGT Gly 255	GAT Asp	768			
GAT Asp	GTA Val	AAT Asn 260	TTT Phe	AAT Asn	GCA Ala	GAA Glu	TTT Phe 265	AAT Asn	AAT Asn	GTA Val	GCA Ala	ATT Ile	ATT Ile 270	TTT Phe	GAT Asp	816			
GAC Asp	TTT Phe	AAA Lys 275	TCA Ser	TCA Ser	CTA Leu	ACA Thr	GGA Gly 280	GGA Gly	GAT Asp	GAC Asp	GGA Gly	TTA Leu 285	ATA Ile	GAT Asp	AAT Asn	864			
GTC Val 290	ATT Ile	GAG Glu	GTT Val	CTT Leu	AAC Asn 295	ACC Thr	GTG Val	AAA Lys	GCA Ala	TTA Leu	CCA Pro 300	TTT Phe	ATA Ile	AAG Lys	AAC Asn	912			
GCC Ala 305	GAC Asp	AGT Ser	AAA Lys	CTA Leu	TAC Tyr 310	AGA Arg	GAA Glu	TTA Leu	GTA Val	ACT Thr 315	AGA Arg	ACA Thr	AAA Lys	GCT Ala	TTA Leu 320	960			
GAG Glu	ACT Thr	CTT Leu	AAA Lys	AAT Asn 325	CAA Gln	ATC Ile	AAA Lys	ACG Thr	ACT Thr 330	GAT Asp	TTG Leu	CCT Pro	CTT Leu	ATA Ile 335	GAT Asp	1008			

- 33 -

GAT Asp	ATA Ile	CCC Pro	GAA Glu 340	ACT Thr	TTG Leu	TCT Ser	CAA Gln	GTG Val 345	AAC Asn	TTT Phe	CCG Pro	AAT Asn	GAC Asp 350	GAA Glu	AAT Asn	1056
CAA Gln	TTG Leu	CCT Pro 355	ACA Thr	CCA Pro	ATA Ile	GGA Gly	AAT Asn 360	TGG Trp	GTT Val	GAT Asp	GGC Gly	GTA Val 365	GAA Glu	GTT Val	AGG Arg	1104
TAC Tyr	GCA Ala 370	GTA Val	CAG Gln	TAT Tyr	GAA Glu	AGT Ser 375	AAG Lys	GGC Gly	ATG Met	TAT Tyr	TCG Ser 380	AAA Lys	TTC Phe	AGT Ser	GAA Glu	1152
TGG Trp 385	TCT Ser	GAA Glu	CCA Pro	TTT Phe	ACT Thr 390	GTC Val	CAA Gln	GGT Gly	AAC Asn	GCT Ala 395	TGT Cys	CCG Pro	ACT Thr	ATA Ile	AAA Lys 400	1200
GTT Val	CGT Arg	GTT Val	GAT Asp	CCG Pro 405	AAA Lys	AAG Lys	AGA Arg	AAT Asn	AGA Arg 410	CTT Leu	ATC Ile	TTT Phe	AGG Arg	AAG Lys 415	TTC Phe	1248
AAC Asn	TCA Ser	GGA Gly	AAA Lys 420	CCT Pro	CAG Gln	TTT Phe	GCT Ala	GGA Gly 425	ACC Thr	ATG Met	ACT Thr	CAT His	TCA Ser 430	CAA Gln	ACA Thr	1296
AAT Asn	TTT Phe	AAA Lys 435	GAT Asp	ATT Ile	CAT His	CGT Arg	GAT Asp 440	CTA Leu	TAC Tyr	GAT Asp	GCA Ala	GCC Ala 445	TTA Leu	AAT Asn	ATT Ile	1344
AAT Asn 450	AAG Lys	TTG Leu	AAA Lys	GCA Ala	GTG Val 455	GAT Glu	GAA Glu	GCT Ala	ACA Thr	ACT Thr	TTG Leu 460	ATT Ile	GAA Glu	AAG Lys	GGT Gly	1392
GCA Ala 465	GAC Asp	ATA Ile	GAA Glu	GCA Ala	AAA Lys 470	TTT Phe	GAC Asp	AAT Asn	GAC Asp	AGA Arg 475	AGT Ser	GCA Ala	ATG Met	CAC His	GCA Ala 480	1440
GTT Val	GCA Ala	TAT Tyr	CGA Arg	GGA Gly 485	AAT Asn	AAC Asn	AAA Lys	ATA Ile	GCC Ala 490	TTA Leu	AGA Arg	TTT Phe	CTT Leu	TTG Leu 495	AAA Lys	1488
AAT Asn	CAA Gln	TCC Ser	ATT Ile 500	GAC Asp	ATC Ile	GAG Glu	TTA Leu	AAA Lys 505	GAT Asp	AAA Lys	AAC Asn	GGC Gly	TTT Phe 510	ACT Thr	CCT Pro	1536
CTA Leu	CAC His	ATC Ile 515	GCA Ala	GCT Ala	GAA Glu	GCA Ala	GGT Gly 520	CAG Gln	GCA Ala	GGA Gly	TTT Phe	GTT Val 525	AAG Lys	TTA Leu	CTA Leu	1584
ATA Ile 530	AAT Asn	CAT His	GGA Gly	GCT Ala	GAT Asp	GTG Val 535	AAT Asn	GCA Ala	AAA Lys	ACA Thr	AGT Ser 540	AAG Lys	ACA Thr	AAT Asn	TTG Leu	1632
ACA Thr 545	CCA Pro	TTA Leu	CAT His	CTT Leu	GCA Ala 550	ACA Thr	CGT Arg	AGT Ser	GGA Gly	TTT Phe 555	TCA Ser	AAA Lys	ACT Thr	GTA Val 560	AGA Arg	1680
AAT Asn	TTA Leu	CTA Leu	GAA Glu	AGC Ser 565	CCA Pro	AAT Asn	ATT Ile	AAG Lys	GTA Val 570	AAT Asn	GAA Glu	AAG Lys	GAG Glu	GAT Asp 575	GAC Asp	1728
GGA Gly	TTT Phe	ACA Thr	CCT Pro 580	TTG Leu	CAT His	ACT Thr	GCA Ala	GTA Val 585	ATG Met	AGT Ser	ACT Thr	TAT Tyr	ATG Met 590	GTT Val	GTC Val	1776
GAT Asp	GCT Ala	TTG Leu	CTA Leu	AAT Asn	CAT His	CCA Pro	GAC Asp	ATT Ile	GAT Asp	AAA Lys	AAT Asn	GCG Ala	CAG Gln	TCT Ser	ACG Thr	1824

595					600 - 34 -					605						
TCA Ser	GGA Gly 610	TTG Leu	ACT Thr	CCT Pro	TTC Phe	CAT His 615	TTA Leu	GCA Ala	ATT Ile	ATT Ile	AAT Asn 620	GAA Glu	AGT Ser	CAA Gln	GAA Glu	1872
GTT Val 625	GCA Ala	GAA Glu	TCT Ser	TTA Leu	GTG Val 630	GAA Glu	AGT Ser	AAT Asn	GCT Ala	GAT Asp 635	CTA Leu	AAT Asn	ATT Ile	CAG Gln	GAT Asp 640	1920
GTT Val	AAC Asn	CAT His	ATG Met	GCT Ala 645	CCT Pro	ATT Ile	CAT His	TTT Phe	GCA Ala 650	GCT Ala	TCA Ser	ATG Met	GGT Gly	AGT Ser 655	ATT Ile	1968
AAA Lys	ATG Met	CTT Leu	AGA Arg 660	TAT Tyr	CTC Leu	ATT Ile	TCC Ser	ATA Ile 665	AAA Lys	GAT Asp	AAA Lys	GTT Val	AGT Ser 670	ATT Ile	AAT Asn	2016
TCT Ser	GTG Val	ACT Thr 675	GAG Glu	AAT Asn	AAT Asn	AAC Asn	TGG Trp 680	ACA Thr	CCT Pro	TTA Leu	CAT His	TTT Phe 685	GCT Ala	ATA Ile	TAT Tyr	2064
TTT Phe 690	AAA Lys	AAA Lys	GAA Glu	GAT Asp	GCT Ala	GCA Ala 695	AAA Lys	GAA Glu	TTG Leu	TTG Leu	AAA Lys 700	CAA Gln	GAT Asp	GAC Asp	ATA Ile	2112
AAT Asn 705	TTA Leu	ACA Thr	ATT Ile	GTT Val	GCA Ala 710	GAT Asp	GGT Gly	AAT Asn	CTT Leu	ACC Thr 715	GTT Val	TTA Leu	CAT His	CTT Leu	GCT Ala 720	2160
GTT Val	TCG Ser	ACA Thr	GGA Gly	CAA Gln 725	ATA Ile	AAT Asn	ATA Ile	ATT Ile	AAA Lys 730	GAA Glu	TTA Leu	TTG Leu	AAG Lys	AGA Arg 735	GGC Gly	2208
TCC Ser	AAT Asn	ATA Ile	GAA Glu 740	GAA Glu	AAA Lys	ACT Thr	GGA Gly	GAA Glu 745	GGA Gly	TAT Tyr	ACA Thr	TCT Ser	CTC Leu 750	CAC His	ATC Ile	2256
GCT Ala	GCG Ala	ATG Met 755	CGA Arg	AAG Lys	GAG Glu	CCA Pro	GAG Glu 760	ATA Ile	GCT Ala	GTT Val	GTT Val	TTG Leu 765	ATT Ile	GAA Glu	AAC Asn	2304
GGT Gly 770	GCT Ala	GAC Asp	ATA Ile	GAA Glu	GCT Ala	CGA Arg 775	TCA Ser	GCT Ala	GAT Asp	AAT Asn	TTA Leu 780	ACA Thr	CCT Pro	TTA Leu	CAT His	2352
TCT Ser 785	GCC Ala	GCA Ala	AAA Lys	ATA Ile	GGA Gly 790	AGG Arg	AAA Lys	TCT Ser	ACA Thr	GTA Val 795	CTT Leu	TAC Tyr	TTA Leu	TTA Leu	GAA Glu 800	2400
AAA Lys	GGA Gly	GCT Ala	GAC Asp	ATT Ile 805	GGA Gly	GCT Ala	AAA Lys	ACA Thr	GCA Ala 810	GAC Asp	GGT Gly	TCT Ser	ACT Thr	GCC Ala 815	TTG Leu	2448
CAT His	TTA Leu	GCT Ala	GTA Val 820	TCT Ser	GGT Gly	CGT Arg	AAA Lys	ATG Met 825	AAA Lys	ACT Thr	GTT Val	GAA Glu	ACT Thr 830	CTA Leu	TTA Leu	2496
AAT Asn	AAA Lys	GGA Gly 835	GCA Ala	AAT Asn	TTA Leu	AAA Lys	GAA Glu 840	TAC Tyr	GAT Asp	AAC Asn	AAT Asn	AAA Lys 845	TAT Tyr	TTG Leu	CCA Pro	2544
ATA Ile	CAT His 850	AAA Lys	GCT Ala	ATT Ile	ATT Ile	AAT Asn 855	GAT Asp	GAC Asp	CTT Leu	GAC Asp	ATG Met 860	GTA Val	CGT Arg	TTG Leu	TTT Phe	2592

- 35 -															
CTT 863	GAA Glu	AAA Lys	GAT Asp	CCC Pro	AGT Ser 870	CTC Leu	AAA Lys	GAT Asp	GAT Asp 875	GAA Glu	ACA Thr	GAA Glu	GAG Glu	GGT Gly	AGA Arg 880
ACT 885	TCA Ser	ATT Ile	ATG Met	TTA Leu 885	ATT Ile	GTT Val	CAG Gln	AAA Lys	TTG Leu 890	CTT Leu	CTT Leu	GAA Glu	TTA Leu	TAT Tyr	AAC Asn 895
TAT 900	TTT Phe	ATA Ile	AAT Asn 900	AAT Asn	TAT Tyr	GCT Ala	GAA Glu	ACT Thr 905	TTG Leu	GAT Asp	GAA Glu	GAA Glu	GCT Ala 910	TTA Leu	TTC Phe
AAC 915	CGC Arg	TTA Leu 915	GAT Asp	GAA Glu	CAA Gln	GGG Gly	AAA Lys 920	TTA Leu	GAG Glu	CTT Leu	GCA Ala	TAT Tyr 925	ATC Ile	TTC Phe	CAT His
AAT 930	AAA Lys	GAA Glu	GGT Gly	GAT Asp	GCA Ala	AAA Lys 935	GAG Glu	GCT Ala	GTT Val	AAG Lys	CCA Pro 940	ACT Thr	ATC Ile	GTT Leu	GTT Val
ACA 945	ATT Ile	AAA Lys	CTT Leu	ATG Met	GAA Glu 950	TAC Tyr	TGC Cys	TTA Leu	AAA Lys 955	AAA Lys	CTT Leu	CGC Arg	GAA Glu	GAG Glu	TCT Ser 960
GGA 965	GCT Ala	CGT Pro	GAA Glu	GGT Gly	AGT Ser	TTC Phe	GAT Asp	TCT Ser	CCA Pro 970	TCT Ser	TCA Ser	AAG Lys	CAA Gln	TGT Cys	ATT Ile 975
TCT 980	ACC Thr	TTT Phe	TCA Ser	GAG Glu	GAT Asp	GAA Glu	ATG Met	TTT Phe 985	CGT Arg	CGT Arg	ACT Thr	TTA Leu	CGC Pro 990	GAA Glu	TGA *

(2) INFORMATION FOR SEQ ID NO: 2:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 992 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Asp Glu Glu Asp Gly Glu Met Thr Leu Glu Glu Arg Gln Ala Gln Cys
 1 5 10 15
 Lys Ala Ile Glu Tyr Ser Asn Ser Val Phe Gly Met Ile Ala Asp Val
 20 25 30
 Ala Asn Asp Ile Gly Ser Ile Pro Val Ile Gly Glu Val Val Gly Ile
 35 40 45
 Val Thr Ala Pro Ile Ala Ile Val Ser His Ile Thr Ser Ala Gly Leu
 50 55 60
 Asp Ile Ala Ser Thr Ala Leu Asp Cys Asp Asp Ile Pro Phe Asp Glu
 65 70 75 80
 Ile Lys Glu Ile Leu Glu Glu Arg Phe Asn Glu Ile Asp Arg Lys Leu
 85 90 95
 Asp Lys Asn Thr Ala Ala Leu Glu Glu Val Ser Lys Leu Val Ser Lys
 100 105 110
 Thr Phe Val Thr Val Glu Lys Thr Arg Asn Glu Met Asn Glu Asn Phe
 115 120 125

- 36 -

Lys Leu Val Leu Glu Thr Ile Glu Ser Lys Glu Ile Lys Ser Ile Val
 130 135 140
 Phe Lys Ile Asn Asp Phe Lys Lys Phe Phe Glu Lys Glu Arg Gln Arg
 145 150 155 160
 Ile Lys Gly Leu Pro Lys Asp Arg Tyr Val Ala Lys Leu Leu Glu Gln
 165 170 175
 Lys Gly Ile Leu Gly Ser Leu Lys Glu Val Arg Glu Pro Ser Gly Asn
 180 185 190
 Ser Leu Ser Ser Ala Leu Asn Glu Leu Leu Asp Lys Asn Asn Asn Tyr
 195 200 205
 Ala Ile Pro Lys Val Val Asp Asp Asn Lys Ala Phe Gln Ala Leu Tyr
 210 215 220
 Ala Leu Phe Tyr Gly Thr Gln Thr Tyr Ala Ala Val Met Phe Phe Leu
 225 230 235 240
 Leu Glu Gln His Ser Tyr Leu Ala Asp Tyr Tyr Tyr Gln Lys Gly Asp
 245 250 255
 Asp Val Asn Phe Asn Ala Glu Phe Asn Asn Val Ala Ile Ile Phe Asp
 260 265 270
 Asp Phe Lys Ser Ser Leu Thr Gly Gly Asp Asp Gly Leu Ile Asp Asn
 275 280 285
 Val Ile Glu Val Leu Asn Thr Val Lys Ala Leu Pro Phe Ile Lys Asn
 290 295 300
 Ala Asp Ser Lys Leu Tyr Arg Glu Leu Val Thr Arg Thr Lys Ala Leu
 305 310 315 320
 Glu Thr Leu Lys Asn Gln Ile Lys Thr Thr Asp Leu Pro Leu Ile Asp
 325 330 335
 Asp Ile Pro Glu Thr Leu Ser Gln Val Asn Phe Pro Asn Asp Glu Asn
 340 345 350
 Gln Leu Pro Thr Pro Ile Gly Asn Trp Val Asp Gly Val Glu Val Arg
 355 360 365
 Tyr Ala Val Gln Tyr Glu Ser Lys Gly Met Tyr Ser Lys Phe Ser Glu
 370 375 380
 Trp Ser Glu Pro Phe Thr Val Gln Gly Asn Ala Cys Pro Thr Ile Lys
 385 390 395 400
 Val Arg Val Asp Pro Lys Lys Arg Asn Arg Leu Ile Phe Arg Lys Phe
 405 410 415
 Asn Ser Gly Lys Pro Gln Phe Ala Gly Thr Met Thr His Ser Gln Thr
 420 425 430
 Asn Phe Lys Asp Ile His Arg Asp Leu Tyr Asp Ala Ala Leu Asn Ile
 435 440 445
 Asn Lys Leu Lys Ala Val Asp Glu Ala Thr Thr Leu Ile Glu Lys Gly
 450 455 460
 Ala Asp Ile Glu Ala Lys Phe Asp Asn Asp Arg Ser Ala Met His Ala
 465 470 475 480

SUBSTITUTE SHEET (RULE 26)

- 37 -

Val Ala Tyr Arg Gly Asn Asn Lys Ile Ala Leu Arg Phe Leu Leu Lys
 485 490 495
 Asn Gln Ser Ile Asp Ile Glu Leu Lys Asp Lys Asn Gly Phe Thr Pro
 500 505 510
 Leu His Ile Ala Ala Glu Ala Gly Gln Ala Gly Phe Val Lys Leu Leu
 515 520 525
 Ile Asn His Gly Ala Asp Val Asn Ala Lys Thr Ser Lys Thr Asn Leu
 530 535 540
 Thr Pro Leu His Leu Ala Thr Arg Ser Gly Phe Ser Lys Thr Val Arg
 545 550 555 560
 Asn Leu Leu Glu Ser Pro Asn Ile Lys Val Asn Glu Lys Glu Asp Asp
 565 570 575
 Gly Phe Thr Pro Leu His Thr Ala Val Met Ser Thr Tyr Met Val Val
 580 585 590
 Asp Ala Leu Leu Asn His Pro Asp Ile Asp Lys Asn Ala Gln Ser Thr
 595 600 605
 Ser Gly Leu Thr Pro Phe His Leu Ala Ile Ile Asn Glu Ser Gln Glu
 610 615 620
 Val Ala Glu Ser Leu Val Glu Ser Asn Ala Asp Leu Asn Ile Gln Asp
 625 630 635 640
 Val Asn His Met Ala Pro Ile His Phe Ala Ala Ser Met Gly Ser Ile
 645 650 655
 Lys Met Leu Arg Tyr Leu Ile Ser Ile Lys Asp Lys Val Ser Ile Asn
 660 665 670
 Ser Val Thr Glu Asn Asn Asn Trp Thr Pro Leu His Phe Ala Ile Tyr
 675 680 685
 Phe Lys Lys Glu Asp Ala Ala Lys Glu Leu Leu Lys Gln Asp Asp Ile
 690 695 700
 Asn Leu Thr Ile Val Ala Asp Gly Asn Leu Thr Val Leu His Leu Ala
 705 710 715 720
 Val Ser Thr Gly Gln Ile Asn Ile Ile Lys Glu Leu Leu Lys Arg Gly
 725 730 735
 Ser Asn Ile Glu Glu Lys Thr Gly Glu Gly Tyr Thr Ser Leu His Ile
 740 745 750
 Ala Ala Met Arg Lys Glu Pro Glu Ile Ala Val Val Leu Ile Glu Asn
 755 760 765
 Gly Ala Asp Ile Glu Ala Arg Ser Ala Asp Asn Leu Thr Pro Leu His
 770 775 780
 Ser Ala Ala Lys Ile Gly Arg Lys Ser Thr Val Leu Tyr Leu Leu Glu
 785 790 795 800
 Lys Gly Ala Asp Ile Gly Ala Lys Thr Ala Asp Gly Ser Thr Ala Leu
 805 810 815
 His Leu Ala Val Ser Gly Arg Lys Met Lys Thr Val Glu Thr Leu Leu
 820 825 830

SUBSTITUTE SHEET (RULE 26)

- 38 -

Asn Lys Gly Ala Asn Leu Lys Glu Tyr Asp Asn Asn Lys Tyr Leu Pro
 835 840 845
 Ile His Lys Ala Ile Ile Asn Asp Asp Leu Asp Met Val Arg Leu Phe
 850 855 860
 Leu Glu Lys Asp Pro Ser Leu Lys Asp Asp Glu Thr Glu Glu Gly Arg
 865 870 875 880
 Thr Ser Ile Met Leu Ile Val Gln Lys Leu Leu Leu Glu Leu Tyr Asn
 885 890 895
 Tyr Phe Ile Asn Asn Tyr Ala Glu Thr Leu Asp Glu Glu Ala Leu Phe
 900 905 910
 Asn Arg Leu Asp Glu Gln Gly Lys Leu Glu Leu Ala Tyr Ile Phe His
 915 920 925
 Asn Lys Glu Gly Asp Ala Lys Glu Ala Val Lys Pro Thr Ile Leu Val
 930 935 940
 Thr Ile Lys Leu Met Glu Tyr Cys Leu Lys Lys Leu Arg Glu Glu Ser
 945 950 955 960
 Gly Ala Pro Glu Gly Ser Phe Asp Ser Pro Ser Ser Lys Gln Cys Ile
 965 970 975
 Ser Thr Phe Ser Glu Asp Glu Met Phe Arg Arg Thr Leu Pro Glu *
 980 985 990

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3706 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PLASMID DNA"

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: LATRODECTUS MACTANS TREDECIMGUTTATUS

(vii) IMMEDIATE SOURCE:

- (B) CLONE: pT7.deltAFL

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 45..3686

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GGTCAATTGA AACTTTATGA TAGGATTCAC TTTCTTATAT AGAA ATG CAT TCC AAA	56
Met His Ser Lys	
995	
GAA TTA CAA ACT ATT TCA GCA GCG GTA GCA CGA AAA GCA GTA CCC AAT	104
Glu Leu Gln Thr Ile Ser Ala Ala Val Ala Arg Lys Ala Val Pro Asn	
1000 1005 1010	
ACT ATG GTT ATT CGG TTG AAA AGA GAT GAA GAA GAT GGA GAA ATG ACT	152
Thr Met Val Ile Arg Leu Lys Arg Asp Glu Glu Asp Gly Glu Met Thr	

- 39 -

1015	1020	1025	
CTA GAA GAA AGA CAA GCA CAA TGC AAA GCA ATA GAG TAC AGC AAT TCA Leu Glu Glu Arg Gln Ala Gln Cys Lys Ala Ile Glu Tyr Ser Asn Ser 1030 1035 1040			200
GTT TTT GGG ATG ATC GCT GAT GTA GCT AAC GAC ATC GGT TCC ATT CCT Val Phe Gly Met Ile Ala Asp Val Ala Asn Asp Ile Gly Ser Ile Pro 1045 1050 1055 1060			248
GTA ATT GGC GAA GTA GTT GGC ATT GTA ACT GCC CCA ATT GCC ATC GTA Val Ile Gly Glu Val Val Gly Ile Val Thr Ala Pro Ile Ala Ile Val 1065 1070 1075			296
AGT CAC ATT ACT AGC GCA GGC TTG GAT ATA GCT TCT ACG GCA TTA GAT Ser His Ile Thr Ser Ala Gly Leu Asp Ile Ala Ser Thr Ala Leu Asp 1080 1085 1090			344
TGT GAT GAT ATA CCT TTT GAT GAG ATT AAG GAA ATA TTA GAA GAA AGA Cys Asp Asp Ile Pro Phe Asp Glu Ile Lys Glu Ile Leu Glu Glu Arg 1095 1100 1105			392
TTC AAT GAA ATA GAT AGA AAG TTG GAC AAG AAC ACA GCT GCT TTG GAA Phe Asn Glu Ile Asp Arg Lys Leu Asp Lys Asn Thr Ala Ala Leu Glu 1110 1115 1120			440
GAG GTC TCT AAA CTG GTA AGT AAA ACT TTT GTT ACG GTG GAA AAA ACA Glu Val Ser Lys Leu Val Ser Lys Thr Phe Val Thr Val Glu Lys Thr 1125 1130 1135 1140			488
AGG AAT GAA ATG AAC GAA AAT TTT AAG CTT GTT TTG GAA ACT ATA GAA Arg Asn Glu Met Asn Glu Asn Phe Lys Leu Val Leu Glu Thr Ile Glu 1145 1150 1155			536
AGC AAA GAA ATA AAA TCA ATT GTA TTC AAA ATA AAT GAT TTT AAA AAG Ser Lys Glu Ile Lys Ser Ile Val Phe Lys Ile Asn Asp Phe Lys Lys 1160 1165 1170			584
TTT TTT GAA AAA GAA CGA CAA AGA ATT AAA GGT TTG CCT AAA GAT AGG Phe Phe Glu Lys Glu Arg Gln Arg Ile Lys Gly Leu Pro Lys Asp Arg 1175 1180 1185			632
TAT GTT GCT AAG CTT CTA GAA CAA AAA GGT ATT TTA GGT TCT TTA AAA Tyr Val Ala Lys Leu Leu Glu Gln Lys Gly Ile Leu Gly Ser Leu Lys 1190 1195 1200			680
GAA GTA AGA GAA CCA TCT GGA AAC AGT CTG AGC TCC GCG TTA AAT GAA Glu Val Arg Glu Pro Ser Gly Asn Ser Leu Ser Ser Ala Leu Asn Glu 1205 1210 1215 1220			728
CTC TTA GAC AAA AAC AAC AAC TAT GCC ATC CCA AAA GTG GTT GAT GAT Leu Leu Asp Lys Asn Asn Asn Tyr Ala Ile Pro Lys Val Val Asp Asp 1225 1230 1235			776
AAT AAG GCC TTT CAG GCG CTG TAT GCT TTA TTT TAT GGA ACT CAG ACT Asn Lys Ala Phe Gln Ala Leu Tyr Ala Leu Phe Tyr Gly Thr Gln Thr 1240 1245 1250			824
TAT GCA GCC GTT ATG TTT TTC TTA CTC GAA CAA CAT TCT TAT CTG GCT Tyr Ala Ala Val Met Phe Phe Leu Leu Glu Gln His Ser Tyr Leu Ala 1255 1260 1265			872
GAT TAT TAT TAC CAA AAA GGT GAT GAT GTA AAT TTT AAT GCA GAA TTT Asp Tyr Tyr Tyr Gln Lys Gly Asp Asp Val Asn Phe Asn Ala Glu Phe 1270 1275 1280			920

- 40 -																
AAT Asn 1285	AAT Asn 1285	GTA Val 1285	GCA Ala 1285	ATT Ile 1290	ATT Ile 1290	TTT Phe 1290	GAT Asp 1290	GAC Asp 1295	TTT Phe 1295	AAA Lys 1295	TCA Ser 1295	TCA Ser 1300	CTA Leu 1300	ACA Thr 1300	GGA Gly 1300	968
GGA Gly 1305	GAT Asp 1305	GAC Asp 1305	GGA Gly 1305	TTA Leu 1305	ATA Ile 1305	GAT Asp 1310	AAT Asn 1310	GTC Val 1310	ATT Ile 1310	GAG Glu 1315	GTT Val 1315	CTT Leu 1315	AAC Asn 1315	ACC Thr 1315	GTG Val 1315	1016
AAA Lys 1320	GCA Ala 1320	TTA Leu 1320	CCA Pro 1320	TTT Phe 1320	ATA Ile 1325	AAG Lys 1325	AAC Asn 1325	GCC Ala 1325	GAC Asp 1330	AGT Ser 1330	AAA Lys 1330	CTA Leu 1330	TAC Tyr 1330	AGA Arg 1330	GAA Glu 1330	1064
TTA Leu 1335	GTA Val 1335	ACT Thr 1335	AGA Arg 1335	ACA Thr 1335	AAA Lys 1340	GCT Ala 1340	TTA Leu 1340	GAG Glu 1340	ACT Thr 1345	CTT Leu 1345	AAA Lys 1345	AAT Asn 1345	CAA Gln 1345	ATC Ile 1345	AAA Lys 1345	1112
ACG Thr 1350	ACT Thr 1350	GAT Asp 1350	TTG Leu 1350	CCT Pro 1355	CTT Leu 1355	ATA Ile 1355	GAT Asp 1355	GAT Asp 1360	ATA Ile 1360	CCC Pro 1360	GAA Glu 1360	ACT Thr 1360	TTG Leu 1360	TCT Ser 1360	CAA Gln 1360	1160
GTG Val 1365	AAC Asn 1365	TTT Phe 1370	CCG Pro 1370	AAT Asn 1370	GAC Asp 1370	GAA Glu 1370	AAT Asn 1375	CAA Gln 1375	TTG Leu 1375	CCT Pro 1375	ACA Thr 1375	CCA Pro 1380	ATA Ile 1380	GGA Gly 1380	AAT Asn 1380	1208
TGG Trp 1385	GTT Val 1385	GAT Asp 1385	GGC Gly 1385	GTA Val 1385	GAA Glu 1385	GTT Val 1390	AGG Arg 1390	TAC Tyr 1390	GCA Ala 1390	GTA Val 1395	CAG Gln 1395	TAT Tyr 1395	GAA Glu 1395	AGT Ser 1395	AAG Lys 1395	1256
GGC Gly 1400	ATG Met 1400	TAT Tyr 1400	TCG Ser 1400	AAA Lys 1405	TTC Phe 1405	AGT Ser 1405	GAA Glu 1405	TGG Trp 1405	TCT Ser 1410	GAA Glu 1410	CCA Pro 1410	TTT Phe 1410	ACT Thr 1410	GTC Val 1410	CAA Gln 1410	1304
GGT Gly 1415	AAC Asn 1415	GCT Ala 1415	TGT Cys 1415	CCG Pro 1420	ACT Thr 1420	ATA Ile 1420	AAA Lys 1420	GTT Val 1420	CGT Arg 1425	GTT Val 1425	GAT Asp 1425	CCG Pro 1425	AAA Lys 1425	AAG Lys 1425	AGA Arg 1425	1352
AAT Asn 1430	AGA Arg 1430	CTT Leu 1430	ATC Ile 1430	TTT Phe 1435	AGG Arg 1435	AAG Lys 1435	TTC Phe 1435	AAC Asn 1440	TCA Ser 1440	GGA Gly 1440	AAA Lys 1440	CCT Pro 1440	CAG Gln 1440	TTT Phe 1440	GCT Ala 1440	1400
GGA Gly 1445	ACC Thr 1445	ATG Met 1450	ACT Thr 1450	CAT His 1450	TCA Ser 1450	CAA Gln 1450	ACA Thr 1455	AAT Asn 1455	TTT Phe 1455	AAA Lys 1455	GAT Asp 1455	ATT Ile 1460	CAT His 1460	CGT Arg 1460	GAT Asp 1460	1448
CTA Leu 1465	TAC Tyr 1465	GAT Asp 1465	GCA Ala 1465	GCC Ala 1465	TTA Leu 1465	AAT Asn 1470	ATT Ile 1470	AAT Asn 1475	AAG Lys 1475	TTG Leu 1475	AAA Lys 1475	GCA Ala 1475	GTG Val 1475	GAT Asp 1475	GAA Glu 1475	1496
GCT Ala 1480	ACA Thr 1480	ACT Thr 1480	TTG Leu 1480	ATT Ile 1485	GAA Glu 1485	AAG Lys 1485	GGT Gly 1485	GCA Ala 1485	GAC Asp 1490	ATA Ile 1490	GAA Glu 1490	GCA Ala 1490	AAA Lys 1490	TTT Phe 1490	GAC Asp 1490	1544
AAT Asn 1495	GAC Asp 1495	AGA Arg 1495	AGT Ser 1495	GCA Ala 1500	ATG Met 1500	CAC His 1500	GCA Ala 1500	GTT Val 1505	GCA Ala 1505	TAT Tyr 1505	CGA Arg 1505	GGA Gly 1505	AAT Asn 1505	AAC Asn 1505	AAA Lys 1505	1592
ATA Ile 1510	GCC Ala 1510	TTA Leu 1510	AGA Arg 1515	TTT Phe 1515	CTT Leu 1515	TTG Leu 1515	AAA Lys 1520	AAT Asn 1520	CAA Gln 1520	TCC Ser 1520	ATT Ile 1520	GAC Asp 1520	ATC Ile 1520	GAG Glu 1520	TTA Leu 1520	1640
AAA Lys 1525	GAT Asp 1525	AAA Lys 1530	AAC Asn 1530	GGC Gly 1530	TTT Phe 1530	ACT Thr 1535	CCT Pro 1535	CTA Leu 1535	CAC His 1535	ATC Ile 1535	GCA Ala 1535	GCT Ala 1540	GAA Glu 1540	GCA Ala 1540	GGT Gly 1540	1688
CAG Gln 1736	GCA Ala 1736	GGA Gly 1736	TTT Phe 1736	GTT Val 1736	AAG Lys 1736	TTA Leu 1736	CTA Leu 1736	ATA Ile 1736	AAT Asn 1736	CAT His 1736	GGA Gly 1736	GCT Ala 1736	GAT Asp 1736	GTG Val 1736	AAT Asn 1736	1736

- 41 -																
1545								1550				1555				
GCA	AAA	ACA	AGT	AAG	ACA	AAT	TTG	ACA	CCA	TTA	CAT	CTT	GCA	ACA	CGT	1724
Ala	Lys	Thr	Ser	Lys	Thr	Asn	Leu	Thr	Pro	Leu	His	Leu	Ala	Thr	Arg	
			1560					1565					1570			
AGT	GGA	TTT	TCA	AAA	ACT	GTA	AGA	AAT	TTA	CTA	GAA	AGC	CCA	AAT	ATT	1832
Ser	Gly	Phe	Ser	Lys	Thr	Val	Arg	Asn	Leu	Leu	Glu	Ser	Pro	Asn	Ile	
		1575					1580					1585				
AAG	GTA	AAT	GAA	AAG	GAG	GAT	GAC	GGA	TTT	ACA	CCT	TTG	CAT	ACT	GCA	1880
Lys	Val	Asn	Glu	Lys	Glu	Asp	Asp	Gly	Phe	Thr	Pro	Leu	His	Thr	Ala	
	1590					1595					1600					
GTA	ATG	AGT	ACT	TAT	ATG	GTT	GTC	GAT	GCT	TTG	CTA	AAT	CAT	CCA	GAC	1928
Val	Met	Ser	Thr	Tyr	Met	Val	Val	Asp	Ala	Leu	Leu	Asn	His	Pro	Asp	
1605					1610					1615					1620	
ATT	GAT	AAA	AAT	GCG	CAG	TCT	ACG	TCA	GGA	TTG	ACT	CCT	TTG	CAT	TTA	1976
Ile	Asp	Lys	Asn	Ala	Gln	Ser	Thr	Ser	Gly	Leu	Thr	Pro	Phe	His	Leu	
				1625					1630					1635		
GCA	ATT	ATT	AAT	GAA	AGT	CAA	GAA	ETT	GCA	GAA	TCT	TTA	GTG	GAA	AGT	2024
Ala	Ile	Ile	Asn	Glu	Ser	Gln	Glu	Val	Ala	Glu	Ser	Leu	Val	Glu	Ser	
			1640					1645					1650			
AAT	GCT	GAT	CTA	AAT	ATT	CAG	GAT	ETT	AAC	CAT	ATG	GCT	CCT	ATT	CAT	2072
Asn	Ala	Asp	Leu	Asn	Ile	Gln	Asp	Val	Asn	His	Met	Ala	Pro	Ile	His	
		1655				1660						1665				
TTT	GCA	GCT	TCA	ATG	GGT	AGT	ATT	AAA	ATG	CTT	AGA	TAT	CTC	ATT	TCC	2120
Phe	Ala	Ala	Ser	Met	Gly	Ser	Ile	Lys	Met	Leu	Arg	Tyr	Leu	Ile	Ser	
	1670					1675					1680					
ATA	AAA	GAT	AAA	GTT	AGT	ATT	AAT	TCT	GTG	ACT	GAG	AAT	AAT	AAC	TGG	2168
Ile	Lys	Asp	Lys	Val	Ser	Ile	Asn	Ser	Val	Thr	Glu	Asn	Asn	Asn	Trp	
1685					1690					1695					1700	
ACA	CCT	TTA	CAT	TTT	GCT	ATA	TAT	TTT	AAA	AAA	GAA	GAT	GCT	GCA	AAA	2216
Thr	Pro	Leu	His	Phe	Ala	Ile	Tyr	Phe	Lys	Lys	Glu	Asp	Ala	Ala	Lys	
				1705					1710					1715		
GAA	TTG	TTG	AAA	CAA	GAT	GAC	ATA	AAT	TTA	ACA	ATT	ETT	GCA	GAT	GGT	2264
Glu	Leu	Leu	Lys	Gln	Asp	Asp	Ile	Asn	Leu	Thr	Ile	Val	Ala	Asp	Gly	
			1720					1725					1730			
AAT	CTT	ACC	GTT	TTA	CAT	CTT	GCT	ETT	TCG	ACA	GGA	CAA	ATA	AAT	ATA	2312
Asn	Leu	Thr	Val	Leu	His	Leu	Ala	Val	Ser	Thr	Gly	Gln	Ile	Asn	Ile	
		1735					1740					1745				
ATT	AAA	GAA	TTA	TTG	AAG	AGA	GGC	TCC	AAT	ATA	GAA	GAA	AAA	ACT	GGA	2360
Ile	Lys	Glu	Leu	Leu	Lys	Arg	Gly	Ser	Asn	Ile	Glu	Glu	Lys	Thr	Gly	
	1750					1755					1760					
GAA	GGA	TAT	ACA	TCT	CTC	CAC	ATC	GCT	GCG	ATG	CGA	AAG	GAG	CCA	GAG	2408
Glu	Gly	Tyr	Thr	Ser	Leu	His	Ile	Ala	Ala	Met	Arg	Lys	Glu	Pro	Glu	
1765					1770					1775					1780	
ATA	GCT	GTT	GTT	TTG	ATT	GAA	AAC	GGT	GCT	GAC	ATA	GAA	GCT	CCA	TCA	2456
Ile	Ala	Val	Val	Leu	Ile	Glu	Asn	Gly	Ala	Asp	Ile	Glu	Ala	Arg	Ser	
				1785				1790						1795		
GCT	GAT	AAT	TTA	ACA	CCT	TTA	CAT	TCT	GCC	GCA	AAA	ATA	GGA	AGG	AAA	2504
Ala	Asp	Asn	Leu	Thr	Pro	Leu	His	Ser	Ala	Ala	Lys	Ile	Gly	Arg	Lys	
			1800					1805					1810			

- 42 -

TCT	ACA	GTA	CTT	TAC	TTA	TTA	GAA	AAA	GGA	GCT	GAC	ATT	GGA	GCT	AAA	2552
Ser	Thr	Val	Leu	Tyr	Leu	Leu	Glu	Lys	Gly	Ala	Asp	Ile	Gly	Ala	Lys	
		1815					1820					1825				
ACA	GCA	GAC	GGT	TCT	ACT	GCC	TTG	CAT	TTA	GCT	GTA	TCT	GGT	CGT	AAA	2600
Thr	Ala	Asp	Gly	Ser	Thr	Ala	Leu	His	Leu	Ala	Val	Ser	Gly	Arg	Lys	
		1830				1835					1840					
ATG	AAA	ACT	GTT	GAA	ACT	CTA	TTA	AAT	AAA	GGA	GCA	AAT	TTA	AAA	GAA	2648
Met	Lys	Thr	Val	Glu	Thr	Leu	Leu	Asn	Lys	Gly	Ala	Asn	Leu	Lys	Glu	
		1845			1850					1855					1860	
TAC	GAT	AAC	AAT	AAA	TAT	TTG	CCA	ATA	CAT	AAA	GCT	ATT	ATT	AAT	GAT	2696
Tyr	Asp	Asn	Asn	Lys	Tyr	Leu	Pro	Ile	His	Lys	Ala	Ile	Ile	Asn	Asp	
				1865					1870					1875		
GAC	CTT	GAC	ATG	GTA	CGT	TTG	TTT	CTT	GAA	AAA	GAT	CCC	AGT	CTC	AAA	2744
Asp	Leu	Asp	Met	Val	Arg	Leu	Phe	Leu	Glu	Lys	Asp	Pro	Ser	Leu	Lys	
			1880					1885					1890			
GAT	GAT	GAA	ACA	GAA	GAG	GGT	AGA	ACT	TCA	ATT	ATG	TTA	ATT	GTT	CAG	2792
Asp	Asp	Glu	Thr	Glu	Glu	Gly	Arg	Thr	Ser	Ile	Met	Leu	Ile	Val	Gln	
		1895				1900						1905				
AAA	TTG	CTT	CTT	GAA	TTA	TAT	AAC	TAT	TTT	ATA	AAT	AAT	TAT	GCT	GAA	2840
Lys	Leu	Leu	Leu	Glu	Leu	Tyr	Asn	Tyr	Phe	Ile	Asn	Asn	Tyr	Ala	Glu	
		1910				1915					1920					
ACT	TTG	GAT	GAA	GAA	GCT	TTA	TTC	AAC	CGC	TTA	GAT	GAA	CAA	GGG	AAA	2888
Thr	Leu	Asp	Glu	Glu	Ala	Leu	Phe	Asn	Arg	Leu	Asp	Glu	Gln	Gly	Lys	
		1925			1930					1935					1940	
TTA	GAG	CTT	GCA	TAT	ATC	TTC	CAT	AAT	AAA	GAA	GGT	GAT	GCA	AAA	GAG	2936
Leu	Glu	Leu	Ala	Tyr	Ile	Phe	His	Asn	Lys	Glu	Gly	Asp	Ala	Lys	Glu	
				1945					1950					1955		
GCT	GTT	AAG	CCA	ACT	ATC	CTT	GTT	ACA	ATT	AAA	CTT	ATG	GAA	TAC	TGC	2984
Ala	Val	Lys	Pro	Thr	Ile	Leu	Val	Thr	Ile	Lys	Leu	Met	Glu	Tyr	Cys	
			1960					1965					1970			
TTA	AAA	AAA	CTT	CGC	GAA	GAG	TCT	GGA	GCT	CCT	GAA	GGT	AGT	TTC	GAT	3032
Leu	Lys	Lys	Leu	Arg	Glu	Glu	Ser	Gly	Ala	Pro	Glu	Gly	Ser	Phe	Asp	
		1975					1980					1985				
TCT	CCA	TCT	TCA	AAG	CAA	TGT	ATT	TCT	ACC	TTT	TCA	GAG	GAT	GAA	ATG	3080
Ser	Pro	Ser	Ser	Lys	Gln	Cys	Ile	Ser	Thr	Phe	Ser	Glu	Asp	Glu	Met	
		1990				1995					2000					
TTT	CGT	CGT	ACT	TTA	CCG	GAA	ATT	GTA	AAA	GAA	ACG	AAC	AGC	AGA	TAT	3128
Phe	Arg	Arg	Thr	Leu	Pro	Glu	Ile	Val	Lys	Glu	Thr	Asn	Ser	Arg	Tyr	
	2005				2010					2015					2020	
TTA	CCA	CTA	AAG	GGC	TTT	TCT	CGC	AGC	CTA	AAT	AAG	TTT	CTC	CCT	TCT	3176
Leu	Pro	Leu	Lys	Gly	Phe	Ser	Arg	Ser	Leu	Asn	Lys	Phe	Leu	Pro	Ser	
				2025					2030					2035		
CTA	AAA	TTT	GCC	GAA	AGT	AAG	AAT	AGC	TAC	AGA	TCT	GAA	AAT	TTT	GTT	3224
Leu	Lys	Phe	Ala	Glu	Ser	Lys	Asn	Ser	Tyr	Arg	Ser	Glu	Asn	Phe	Val	
			2040					2045					2050			
AGC	AAT	ATT	GAT	TCC	AAC	GGA	GCA	TTA	CTT	TTA	CTC	GAT	GTA	TTT	ATC	3272
Ser	Asn	Ile	Asp	Ser	Asn	Gly	Ala	Leu	Leu	Leu	Leu	Asp	Val	Phe	Ile	
		2055				2060						2065				
AGA	AAG	TTT	ACT	AAT	GAG	AAA	TAC	AAT	TTG	ACT	GGA	AAA	GAA	GCT	GTA	3320
Arg	Lys	Phe	Thr	Asn	Glu	Lys	Tyr	Asn	Leu	Thr	Gly	Lys	Glu	Ala	Val	

SUBSTITUTE SHEET (RULE 26)

- 43 -

2070	2075	2080	
CCC TAT CTG GAA GCA AAG GCT TCA TCA TTA CGT ATC GCT TCT AAA TTT Pro Tyr Leu Glu Ala Lys Ala Ser Ser Leu Arg Ile Ala Ser Lys Phe 2085 2090 2095 2100			3368
GAA GAA CTT CTA ACT GAA GTT AAA GGT ATT CCG GCT GGA GAG CTA ATT Glu Glu Leu Leu Thr Glu Val Lys Gly Ile Pro Ala Gly Glu Leu Ile 2105 2110 2115			3416
AAT ATG GCC GAA GTG AGT TCC AAC ATA CAT AAG GCA ATT GCA AGT GGT Asn Met Ala Glu Val Ser Ser Asn Ile His Lys Ala Ile Ala Ser Gly 2120 2125 2130			3464
AAG CCT GTA TCA AAA GTC TTA TGT TCG TAT TTG GAT ACC TTT TCT GAA Lys Pro Val Ser Lys Val Leu Cys Ser Tyr Leu Asp Thr Phe Ser Glu 2135 2140 2145			3512
TTA AAT TCT CAA CAA ATG GAA GAA TTA GTT AAC ACA TAC TTA TCC ACC Leu Asn Ser Gln Gln Met Glu Glu Leu Val Asn Thr Tyr Leu Ser Thr 2150 2155 2160			3560
AAA CCT TCT GTA ATT ACG TCA GCA TCT GCA GAT TAC CAG AAA CTT CCT Lys Pro Ser Val Ile Thr Ser Ala Ser Ala Asp Tyr Gln Lys Leu Pro 2165 2170 2175 2180			3608
AAT TTG TTA ACT GCA ACT TGC TTA GAA CCA GAA AGA ATG GCT CAA CTT Asn Leu Leu Thr Ala Thr Cys Leu Glu Pro Glu Arg Met Ala Gln Leu 2185 2190 2195			3656
ATA GAT GTG CAT CAA AAG ATG TTT TTA CGT TAAAATACCA TTCCTTCTGT Ile Asp Val His Gln Lys Met Phe Leu Arg 2200 2205			3706

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1214 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met His Ser Lys Glu Leu Gln Thr Ile Ser Ala Ala Val Ala Arg Lys
1 5 10 15

Ala Val Pro Asn Thr Met Val Ile Arg Leu Lys Arg Asp Glu Glu Asp
20 25 30

Gly Glu Met Thr Leu Glu Glu Arg Gln Ala Gln Cys Lys Ala Ile Glu
35 40 45

Tyr Ser Asn Ser Val Phe Gly Met Ile Ala Asp Val Ala Asn Asp Ile
50 55 60

Gly Ser Ile Pro Val Ile Gly Glu Val Val Gly Ile Val Thr Ala Pro
65 70 75 80

Ile Ala Ile Val Ser His Ile Thr Ser Ala Gly Leu Asp Ile Ala Ser
85 90 95

Thr Ala Leu Asp Cys Asp Asp Ile Pro Phe Asp Glu Ile Lys Glu Ile
100 105 110

- 44 -

Leu Glu Glu Arg Phe Asn Glu Ile Asp Arg Lys Leu Asp Lys Asn Thr
 115 120 125
 Ala Ala Leu Glu Glu Val Ser Lys Leu Val Ser Lys Thr Phe Val Thr
 130 135 140
 Val Glu Lys Thr Arg Asn Glu Met Asn Glu Asn Phe Lys Leu Val Leu
 145 150 155 160
 Glu Thr Ile Glu Ser Lys Glu Ile Lys Ser Ile Val Phe Lys Ile Asn
 165 170 175
 Asp Phe Lys Lys Phe Phe Glu Lys Glu Arg Gln Arg Ile Lys Gly Leu
 180 185 190
 Pro Lys Asp Arg Tyr Val Ala Lys Leu Leu Glu Gln Lys Gly Ile Leu
 195 200 205
 Gly Ser Leu Lys Glu Val Arg Glu Pro Ser Gly Asn Ser Leu Ser Ser
 210 215 220
 Ala Leu Asn Glu Leu Leu Asp Lys Asn Asn Asn Tyr Ala Ile Pro Lys
 225 230 235 240
 Val Val Asp Asp Asn Lys Ala Phe Gln Ala Leu Tyr Ala Leu Phe Tyr
 245 250 255
 Gly Thr Gln Thr Tyr Ala Ala Val Met Phe Phe Leu Leu Glu Gln His
 260 265 270
 Ser Tyr Leu Ala Asp Tyr Tyr Tyr Gln Lys Gly Asp Asp Val Asn Phe
 275 280 285
 Asn Ala Glu Phe Asn Asn Val Ala Ile Ile Phe Asp Asp Phe Lys Ser
 290 295 300
 Ser Leu Thr Gly Gly Asp Asp Gly Leu Ile Asp Asn Val Ile Glu Val
 305 310 315 320
 Leu Asn Thr Val Lys Ala Leu Pro Phe Ile Lys Asn Ala Asp Ser Lys
 325 330 335
 Leu Tyr Arg Glu Leu Val Thr Arg Thr Lys Ala Leu Glu Thr Leu Lys
 340 345 350
 Asn Gln Ile Lys Thr Thr Asp Leu Pro Leu Ile Asp Asp Ile Pro Glu
 355 360 365
 Thr Leu Ser Gln Val Asn Phe Pro Asn Asp Glu Asn Gln Leu Pro Thr
 370 375 380
 Pro Ile Gly Asn Trp Val Asp Gly Val Glu Val Arg Tyr Ala Val Gln
 385 390 395 400
 Tyr Glu Ser Lys Gly Met Tyr Ser Lys Phe Ser Glu Trp Ser Glu Pro
 405 410 415
 Phe Thr Val Gln Gly Asn Ala Cys Pro Thr Ile Lys Val Arg Val Asp
 420 425 430
 Pro Lys Lys Arg Asn Arg Leu Ile Phe Arg Lys Phe Asn Ser Gly Lys
 435 440 445
 Pro Gln Phe Ala Gly Thr Met Thr His Ser Gln Thr Asn Phe Lys Asp
 450 455 460

- 45 -

Ile 465	His	Arg	Asp	Leu	Tyr 470	Asp	Ala	Ala	Leu	Asn 475	Ile	Asn	Lys	Leu	Lys 480
Ala	Val	Asp	Glu	Ala 485	Thr	Thr	Leu	Ile	Glu 490	Lys	Gly	Ala	Asp	Ile	Glu 495
Ala	Lys	Phe	Asp 500	Asn	Asp	Arg	Ser	Ala 505	Met	His	Ala	Val	Ala 510	Tyr	Arg
Gly	Asn	Asn 515	Lys	Ile	Ala	Leu	Arg 520	Phe	Leu	Leu	Lys	Asn 525	Gln	Ser	Ile
Asp	Ile 530	Glu	Leu	Lys	Asp	Lys 535	Asn	Gly	Phe	Thr	Pro 540	Leu	His	Ile	Ala
Ala 545	Glu	Ala	Gly	Gln	Ala 550	Gly	Phe	Val	Lys	Leu 555	Leu	Ile	Asn	His	Gly 560
Ala	Asp	Val	Asn	Ala 565	Lys	Thr	Ser	Lys	Thr 570	Asn	Leu	Thr	Pro	Leu	His 575
Leu	Ala	Thr	Arg 580	Ser	Gly	Phe	Ser	Lys 585	Thr	Val	Arg	Asn	Leu 590	Leu	Glu
Ser	Pro	Asn 595	Ile	Lys	Val	Asn	Glu 600	Lys	Glu	Asp	Asp	Gly 605	Phe	Thr	Pro
Leu 610	His	Thr	Ala	Val	Met	Ser 615	Thr	Tyr	Met	Val	Val 620	Asp	Ala	Leu	Leu
Asn 625	His	Pro	Asp	Ile	Asp 630	Lys	Asn	Ala	Gln	Ser 635	Thr	Ser	Gly	Leu	Thr 640
Pro	Phe	His	Leu	Ala 645	Ile	Ile	Asn	Glu	Ser 650	Gln	Glu	Val	Ala	Glu	Ser 655
Leu	Val	Glu	Ser 660	Asn	Ala	Asp	Leu	Asn 665	Ile	Gln	Asp	Val	Asn 670	His	Met
Ala	Pro	Ile 675	His	Phe	Ala	Ala	Ser 680	Met	Gly	Ser	Ile	Lys 685	Met	Leu	Arg
Tyr 690	Leu	Ile	Ser	Ile	Lys	Asp 695	Lys	Val	Ser	Ile	Asn 700	Ser	Val	Thr	Glu
Asn 705	Asn	Asn	Trp	Thr	Pro 710	Leu	His	Phe	Ala	Ile 715	Tyr	Phe	Lys	Lys	Glu 720
Asp	Ala	Ala	Lys	Glu 725	Leu	Leu	Lys	Gln	Asp 730	Asp	Ile	Asn	Leu	Thr	Ile 735
Val	Ala	Asp	Gly 740	Asn	Leu	Thr	Val	Leu 745	His	Leu	Ala	Val	Ser 750	Thr	Gly
Gln	Ile	Asn 755	Ile	Ile	Lys	Glu	Leu 760	Leu	Lys	Arg	Gly	Ser 765	Asn	Ile	Glu
Glu 770	Lys	Thr	Gly	Glu	Gly	Tyr 775	Thr	Ser	Leu	His	Ile 780	Ala	Ala	Met	Arg
Lys 785	Glu	Pro	Glu	Ile	Ala 790	Val	Val	Leu	Ile	Glu 795	Asn	Gly	Ala	Asp	Ile 800
Glu	Ala	Arg	Ser	Ala 805	Asp	Asn	Leu	Thr	Pro 810	Leu	His	Ser	Ala	Ala 815	Lys

- 46 -

Ile Gly Arg Lys Ser Thr Val Leu Tyr Leu Leu Glu Lys Gly Ala Asp
 820 825 830
 Ile Gly Ala Lys Thr Ala Asp Gly Ser Thr Ala Leu His Leu Ala Val
 835 840 845
 Ser Gly Arg Lys Met Lys Thr Val Glu Thr Leu Leu Asn Lys Gly Ala
 850 855 860
 Asn Leu Lys Glu Tyr Asp Asn Asn Lys Tyr Leu Pro Ile His Lys Ala
 865 870 875 880
 Ile Ile Asn Asp Asp Leu Asp Met Val Arg Leu Phe Leu Glu Lys Asp
 885 890 895
 Pro Ser Leu Lys Asp Asp Glu Thr Glu Glu Gly Arg Thr Ser Ile Met
 900 905 910
 Leu Ile Val Gln Lys Leu Leu Leu Glu Leu Tyr Asn Tyr Phe Ile Asn
 915 920 925
 Asn Tyr Ala Glu Thr Leu Asp Glu Glu Ala Leu Phe Asn Arg Leu Asp
 930 935 940
 Glu Gln Gly Lys Leu Glu Leu Ala Tyr Ile Phe His Asn Lys Glu Gly
 945 950 955 960
 Asp Ala Lys Glu Ala Val Lys Pro Thr Ile Leu Val Thr Ile Lys Leu
 965 970 975
 Met Glu Tyr Cys Leu Lys Lys Leu Arg Glu Glu Ser Gly Ala Pro Glu
 980 985 990
 Gly Ser Phe Asp Ser Pro Ser Ser Lys Gln Cys Ile Ser Thr Phe Ser
 995 1000 1005
 Glu Asp Glu Met Phe Arg Arg Thr Leu Pro Glu Ile Val Lys Glu Thr
 1010 1015 1020
 Asn Ser Arg Tyr Leu Pro Leu Lys Gly Phe Ser Arg Ser Leu Asn Lys
 1025 1030 1035 1040
 Phe Leu Pro Ser Leu Lys Phe Ala Glu Ser Lys Asn Ser Tyr Arg Ser
 1045 1050 1055
 Glu Asn Phe Val Ser Asn Ile Asp Ser Asn Gly Ala Leu Leu Leu Leu
 1060 1065 1070
 Asp Val Phe Ile Arg Lys Phe Thr Asn Glu Lys Tyr Asn Leu Thr Gly
 1075 1080 1085
 Lys Glu Ala Val Pro Tyr Leu Glu Ala Lys Ala Ser Ser Leu Arg Ile
 1090 1095 1100
 Ala Ser Lys Phe Glu Glu Leu Leu Thr Glu Val Lys Gly Ile Pro Ala
 1105 1110 1115 1120
 Gly Glu Leu Ile Asn Met Ala Glu Val Ser Ser Asn Ile His Lys Ala
 1125 1130 1135
 Ile Ala Ser Gly Lys Pro Val Ser Lys Val Leu Cys Ser Tyr Leu Asp
 1140 1145 1150
 Thr Phe Ser Glu Leu Asn Ser Gln Gln Met Glu Glu Leu Val Asn Thr
 1155 1160 1165

- 47 -

Tyr Leu Ser Thr Lys Pro Ser Val Ile Thr Ser Ala Ser Ala Asp Tyr
1170 1175 1180

Gln Lys Leu Pro Asn Leu Leu Thr Ala Thr Cys Leu Glu Pro Glu Arg
1185 1190 1195 1200

Met Ala Gln Leu Ile Asp Val His Gln Lys Met Phe Leu Arg
1205 1210

- 48 -

CLAIMS

1. A polypeptide, such as a toxin, formed by expression of a truncated form of a gene sequence, or an analogue thereof.
2. A polypeptide as claimed in claim 1, in which the polypeptide is a neurotoxin.
3. A polypeptide as claimed in any preceding claim, in which the polypeptide corresponds to a toxic derivative of a substantially non-toxic precursor polypeptide encoded by the gene sequence.
4. A polypeptide as claimed in any preceding claim, in which the polypeptide comprises an amino acid sequence that corresponds to a truncated form of the amino acid sequence of a substantially non-toxic precursor polypeptide.
5. A polypeptide as claimed in claim 4, in which the amino acid sequence of the polypeptide corresponds to the amino acid sequence of the precursor polypeptide with truncation thereof principally at the carboxy (C) end.
6. A polypeptide as claimed in claim 5, in which truncation is by about 150 to 200 amino acids.

- 49 -

7. A polypeptide as claimed in any of claims 4 to 6, in which the polypeptide amino acid sequence in addition corresponds to the precursor polypeptide amino acid sequence truncated at the amino end (N).
8. A polypeptide as claimed in claim 7, in which the truncation is by less than 50 amino acids, and desirably by 7 or 28 amino acids.
9. A polypeptide as claimed in any preceding claim, in which the amino acid sequence of the polypeptide is homologous to the amino acid sequence of the insect specific neurotoxin δ -Latroinsectotoxin (δ -LIT) or an active derivative thereof.
10. A polypeptide as claimed in any preceding claim, in which the polypeptide comprises an amino acid sequence as shown in SEQIDN01 and SEQIDN02 or an active derivative thereof.
11. A polypeptide as claimed in any preceding claim, in which the toxin is expressed from a nucleotide construct or truncated form of a gene sequence comprising a sequence as shown in SEQIDN01, or active variants thereof.
12. A polypeptide as claimed in any preceding claim, in

- 50 -

which the polypeptide is expressed from a sequence substantially as provided in a microorganism deposited at The National Collections of Industrial and Marine Bacteria Limited, under Accession No. NCIMB 40632.

13. A protein for use as a toxin comprising an amino acid sequence substantially as shown in SEQIDN01 and SEQIDN02, or an active derivative thereof.

14. A nucleotide sequence comprising a truncated form of a gene sequence or an analogue thereof, for use in the expression of a polypeptide, such as a toxin.

15. A nucleotide sequence as claimed in claim 14, in which the nucleotide sequence corresponds to a gene encoding for a precursor polypeptide and truncated at the 3' end thereof, or an active derivative thereof.

16. A nucleotide sequence as claimed in claim 15, in which the nucleotide sequence corresponds to the gene truncated by about 400 to 650 nucleotide bases, and desirably between 550 to 600 nucleotide bases.

17. A nucleotide sequence as claimed in any of claims 14 to 16, in which the nucleotide sequence corresponds to the gene truncated at the 5' thereof.

- 51 -

18. A nucleotide sequence as claimed in claim 17, in which the truncation is by less than 100 nucleotide bases, and desirably by either 84 or 21 nucleotide bases.

19. A nucleotide sequence as claimed in any of claims 14 to 18, in which the nucleotide sequence corresponds to part of a gene encoding for a neurotoxin in the venom of the Black Widow Spider (*Latrodectus mactans* *Tredecimguttatus*), or an active derivative thereof.

20. A nucleotide sequence as claimed in claim 19, in which the nucleotide sequence corresponds to part of the gene encoding the precursor polypeptide of insect specific toxin δ -Lactoinsectotoxin (δ -LIT), or an active derivative thereof.

21. A nucleotide sequence as claimed in any of claims 14 to 20, in which the nucleotide sequence codes for a polypeptide comprising a sequence of 991 amino acids.

22. A nucleotide sequence as claimed in any of claims 14 to 21, in which the nucleotide sequence comprises a base sequence as shown in SEQIDN01, or an active derivative thereof.

23. A nucleotide sequence as claimed in any of claims 14 to 22, in which the nucleotide sequences comprises a

- 52 -

base sequence substantially as comprised in a micro-organism deposited under Accession No. NCIMB 40632 at The National Collections of Industrial and Marine Bacteria Limited.

24. A nucleotide sequence as claimed in any of claims 14 to 23, in which the nucleotide sequence codes for a polypeptide having an amino acid sequence as shown in SEQIDN01 and SEQIDN02, or an active derivative thereof.

25. A nucleotide sequence as claimed in any of claims 14 to 24, in which the nucleotide sequence is a cDNA derived from mRNA by the use of an enzyme such as reverse transcriptase.

26. A nucleotide sequence as claimed in any of claims 14 to 25, in which the nucleotide sequence is an oligonucleotide DNA construct produced perhaps using the polymerase chain reaction (PCR).

27. A method of producing a polypeptide, the method comprising producing a recombinant DNA molecule comprising a truncated form of a gene, and expressing the truncated form in a host expression system, such as a viral or bacterial expression system, to produce the polypeptide.

28. A method as claimed in claim 27, in which the

- 53 -

polypeptide produced is an active toxin substantially as claimed in any preceding claim.

29. A method as claimed in claim 27 or claim 28, in which the truncated form comprises part of a gene which encodes for a non-toxic precursor polypeptide.

30. A method as claimed in any of claims 27 to 29, in which the truncated form comprises a nucleotide sequence substantially as claimed in any of claims 14 to 26.

31. A method as claimed in any of claims 27 to 30, in which the expression system comprises E.coli BL21 (DE3) bacterial cells transformed with pT7-7 vectors comprising the truncated form of the sequence.

32. A method as claimed in any of claims 27 to 31, in which the expression system comprises a baculovirus system.

33. A recombinant DNA molecule comprising a truncated form of a gene encoding for a toxin generally as claimed in any preceding claim.

34. A recombinant DNA molecule as claimed in claim 33, in which the molecule comprises a virus.

35. A recombinant DNA molecule as claimed in claim 34,

- 54 -

in which the molecule comprises a baculovirus.

36. A recombinant DNA molecule substantially as provided in the microorganism deposited under Accession No. NCIMB 40632.

37. An expression vector comprising a truncated form of a gene generally as claimed in any of claims 14 to 26.

38. A cell, such as a viral or bacterial cell transformed with a recombinant molecule substantially as claimed in any of claims 33 to 37.

39. An insecticide comprising a toxin substantially as claimed in any of claims 1 to 13.

40. An insecticide as claimed in claim 39, in which the insecticide is so as to be administered orally or topically.

41. An insecticide as claimed in claim 39 or claim 40, in which the insecticide comprises a spray.

42. An insecticide system comprising means for expressing a truncated form of a gene to produce a toxin substantially as claimed in any preceding claim in an insect to kill or incapacitate the insect.

- 55 -

43. An insecticide system as claimed in claim 42, in which the insecticide system comprises a viral expression system.

44. An insecticide system as claimed in claim 43, in which the viral expression system comprises a baculovirus expression system.

45. A plant comprising a genetically modified cell containing a truncated form of a gene sequence substantially as claimed in any of claims 14 to 26.

46. A non-human animal comprising a genetically modified cell containing a truncated form of a gene sequence substantially as claimed in any of claims 14 to 26.

47. A toxin formed by processing of a substantially isolated non-toxic precursor polypeptide.

48. A toxin as claimed in claim 47, in which the toxin is formed by truncation toward the carboxy (C) end of the precursor polypeptide.

49. A toxin as claimed in claim 48, in which the toxin amino acid sequence generally corresponds to the amino acid sequence of the precursor polypeptide, truncated by between 150 and 200 amino acids.

- 56 -

50. A toxin as claimed in any of claims 47 to 49, in which the toxin amino acid sequence is formed by truncation toward the amino (N) end of the precursor polypeptide amino acid sequence.

51. A toxin as claimed in claim 50, in which the fragment cleaved from the amino end is significantly smaller than the fragment cleaved from the carboxy end.

52. A toxin as claimed in claim 50 or claim 51, in which the fragment cleaved off comprises 7 or 28 amino acids.

53. A toxin as claimed in any of claims 47 to 52, in which the toxin has an amino acid sequence corresponding to a polypeptide encoded by part of a gene of the Black Widow Spider (*Latrodectus mactans* *Tredecimguttatus*).

54. A toxin as claimed in any of claims 47 to 53, in which the toxin comprises or is an analogue of the insect specific neurotoxin δ -Latroinsectotoxin (δ -LIT), or an active derivative thereof.

55. A toxin as claimed in any of claims 47 to 54, in which the toxin comprises an amino acid sequence as shown in SEQIDN01 and SEQIDN02 or an active derivative thereof.

56. A method of producing an active polypeptide from an

- 57 -

isolated inactive precursor polypeptide, the method comprising truncating the isolated precursor polypeptide.

57. A method as claimed in claim 56, in which the isolated precursor polypeptide is truncated at the Carboxyl end.

58. A method as claimed in claim 56 or claim 57, in which the truncation is effected using proteolytic cleavage, and preferably by site directed mutagenesis.

59. A method as claimed in any of claims 56 to 58, in which truncation of the N terminus may be provided.

60. A method as claimed in claims 56 to 59, in which the active polypeptide is a toxin and is substantially as claimed in any of claims 1 to 13, 47 to 55.

61. An isolated nucleotide base sequence encoding for a toxin precursor polypeptide with an amino acid sequence as shown in SEQIDN04 or a derivative thereof.

62. An isolated base sequence comprising a base sequence as shown in SEQIDN03 or a derivative thereof.

63. An isolated base sequence as claimed in any of claims 61 or 62, in which the nucleotide base sequence encodes a precursor polypeptide of the neurotoxin

- 58 -

δ-Latroinsectotoxin (δ-LIT).

64. An isolated base sequence substantially as provided in the microorganism deposited under Accession No. NCIMB 40633.

65. A recombinant DNA molecule comprising a sequence substantially as claimed in any of claims 61 to 64.

66. A recombinant molecule as claimed in claim 65, in which the molecule comprises a virus.

67. A recombinant molecule as claimed in claim 66, in which the virus comprises a baculovirus.

68. A cell, such as a bacterial or viral cell, transformed with a recombinant DNA molecule substantially as claimed in any of claims 65 to 67.

69. An insecticide system comprising means for expressing a base sequence substantially as claimed in any of claims 61 to 64 to produce a precursor polypeptide and to process the precursor polypeptide to produce a toxin in an insect to kill or incapacitate the insect.

70. An insecticide system as claimed in claim 69, in which the system comprises a viral expression system.

- 59 -

71. An insecticide system as claimed in claim 69, in which the viral expression system comprises baculovirus.
72. A plant comprising a genetically modified cell containing a nucleotide sequence substantially as claimed in any of claims 61 to 64.
73. A non-human animal comprising a genetically modified cell containing a nucleotide sequence substantially as claimed in any of claims 61 to 64.
74. A novel toxin substantially as hereinbefore described with reference to SEQIDN01 and SEQIDN02.
75. A nucleotide sequence substantially as hereinbefore described with reference to SEQIDN01.
76. An isolated polypeptide substantially as hereinbefore described with reference to SEQIDN03 and SEQIDN04.
77. An isolated nucleotide sequence substantially as hereinbefore described with reference to SEQIDN03.
78. Any novel subject matter or combination including novel subject matter disclosed, whether or not within the scope of or relating to the same invention as any of the preceding claims.

AMENDED CLAIMS

[received by the International Bureau on 14 September 1995 (14.09.95);
original claims 1-78 replaced by amended claims 1-74 (9 pages)]

1. An active insect specific neurotoxin polypeptide, formed by expression of a truncated form of a gene sequence corresponding to part of a gene encoding for an insect specific neurotoxin present in the venom of the Black Widow Spider (*Latrodectus mactans*
5 *Tredecimguttatus*), or an analogue thereof.
2. A polypeptide as claimed in any preceding claim, in which the polypeptide corresponds to a toxic derivative of a substantially non-toxic precursor polypeptide encoded by the gene sequence.
3. A polypeptide as claimed in any preceding claim, in which the polypeptide
10 comprises an amino acid sequence that corresponds to a truncated form of the amino acid sequence of a substantially non-toxic precursor polypeptide.
4. A polypeptide as claimed in claim 3, in which the amino acid sequence of the polypeptide corresponds to the amino acid sequence of the precursor polypeptide with truncation thereof principally at the carboxy (C) end.
- 15 5. A polypeptide as claimed in claim 4, in which truncation is by about 150 to 200 amino acids.
6. A polypeptide as claimed in any of claims 3 to 5, in which the polypeptide amino acid sequence in addition corresponds to the precursor polypeptide amino acid sequence truncated at the amino end (N).
- 20 7. A polypeptide as claimed in claim 6, in which the truncation is by less than 50 amino acids and desirably by 7 or 28 amino acids.
8. A polypeptide as claimed in any preceding claim, in which the amino acid sequence of the polypeptide is homologous to the amino acid sequence of the insect specific neurotoxin δ -Latroinsectotoxin (δ -LIT) or an active derivative thereof.

9. A polypeptide as claimed in any preceding claim, in which the polypeptide comprises an amino acid sequence as shown in SEQIDN01 and SEQIDN02 or an active derivative thereof.
10. A polypeptide as claimed in any preceding claim, in which the toxin is expressed
5 from a nucleotide construct or truncated form of a gene sequence comprising a sequence as shown in SEQIDN01, or active variants thereof.
11. A polypeptide as claimed in any preceding claim, in which the polypeptide is expressed from a sequence substantially as provided in a microorganism deposited at The National Collections of Industrial and Marine Bacteria Limited, under Accession
10 No. NCIMB 40632.
12. A protein for use as a toxin comprising an amino acid sequence substantially as shown in SEQIDN01 and SEQIDN02, or an active derivative thereof.
13. A nucleotide sequence comprising a truncated form of a gene sequence corresponding to part of a gene encoding for an insect specific neurotoxin present in the
15 venom of the Black Widow Spider (*Latrodectus mactans Tredecimguttatus*) or an analogue thereof, for use in the expression of an active insect specific neurotoxin polypeptide.
14. A nucleotide sequence as claimed in claim 13, in which the nucleotide sequence corresponds to a gene encoding for a precursor polypeptide and truncated at the 3' end thereof, or an active derivative thereof.
- 20 15. A nucleotide sequence as claimed in claim 14, in which the nucleotide sequence corresponds to the gene truncated by about 400 to 650 nucleotide bases, and desirably between 550 to 600 nucleotide bases.

16. A nucleotide sequence as claimed in any of claims 13 to 15, in which the nucleotide sequence corresponds to the gene truncated at the 5' thereof.
17. A nucleotide sequence as claimed in claim 16, in which the truncation is by less than 100 nucleotide bases, and desirably by either 84 or 21 nucleotide bases.
- 5 18. A nucleotide sequence as claimed in any one of claims 13 to 17, in which the nucleotide sequence corresponds to part of the gene encoding the precursor polypeptide of insect specific toxin δ -Lactoinsectotoxin (δ -LIT), or an active derivative thereof.
19. A nucleotide sequence as claimed in any of claims 13 to 18, in which the nucleotide sequence codes for a polypeptide comprising a sequence of 991 amino acids.
- 10 20. A nucleotide sequence as claimed in any of claims 13 to 19, in which the nucleotide sequence comprises a base sequence as shown in SEQIDN01, or an active derivative thereof.
21. A nucleotide sequence as claimed in any of claims 13 to 20, in which the nucleotide sequence comprises a base sequence substantially as comprised in a microorganism
15 deposited under Accession No. NCIMB 40632 at The National Collections of Industrial and Marine Bacteria Limited.
22. A nucleotide sequence as claimed in any of claims 13 to 21, in which the nucleotide sequence codes for a polypeptide having an amino acid sequence as shown in SEQIDN01 and SEQIDN02, or an active derivative thereof.
- 20 23. A nucleotide sequence as claimed in any of claims 13 to 22, in which the nucleotide sequence is a cDNA derived from mRNA by the use of an enzyme such as reverse transcriptase.

24. A nucleotide sequence as claimed in any of claims 13 to 23, in which the nucleotide sequence is an oligonucleotide DNA construct produced perhaps using the polymerase chain reaction (PCR).
25. A method of producing a polypeptide as claimed in claim 1, the method comprising
5 producing a recombinant DNA molecule comprising a truncated form of a gene as claimed in claim 13, and expressing the truncated form in a host expression system, such as a bacterial expression system, to produce the polypeptide.
26. A method as claimed in claim 25, in which the polypeptide produced is an active toxin substantially as claimed in any preceding claim.
- 10 27. A method as claimed in claim 25 or claim 26, in which the truncated form comprises part of a gene which encodes for a non-toxic precursor polypeptide.
28. A method as claimed in any of claims 25 to 27, in which the truncated form comprises a nucleotide sequence substantially as claimed in any of claims 14 to 24.
29. A method as claimed in any of claims 25 to 28, in which the expression system
15 comprises *E. coli* BL21 (DE3) bacterial cells transformed with pT7-7 vectors comprising the truncated form of the sequence.
30. A method as claimed in any of claims 25 to 29, in which the expression system comprises a baculovirus system.
31. A recombinant DNA molecule comprising a truncated form of a gene encoding for
20 a toxin generally as claimed in any preceding claim.
32. A recombinant DNA molecule as claimed in claim 31, in which the molecule comprises a virus.

33. A recombinant DNA molecule as claimed in claim 32, in which the molecule comprises a baculovirus.
34. A recombinant DNA molecule substantially as provided in the microorganism deposited under Accession No. NCIMB 40632.
- 5 35. An expression vector comprising a truncated form of a gene generally as claimed in any of claims 14 to 24.
36. A cell, such as a viral or bacterial cell transformed with a recombinant molecule substantially as claimed in any of claims 31 or 35.
37. An insecticide comprising a toxin substantially as claimed in any of claims 1 to 12.
- 10 38. An insecticide as claimed in claim 37, in which the insecticide is so as to be administered orally or topically.
39. An insecticide as claimed in claim 37 or claim 38, in which the insecticide comprises a spray.
40. An insecticide system comprising means for expressing a truncated form of a gene
15 as claimed in any one of claims 13 to 24 to produce a toxin substantially as claimed in any preceding claim in an insect to kill or incapacitate the insect.
41. An insecticide system as claimed in claim 40, in which the insecticide system comprises a viral expression system.
42. An insecticide system as claimed in claim 41, in which the viral expression system
20 comprises a baculovirus expression system.

43. A plant comprising a genetically modified cell containing a truncated form of a gene sequence substantially as claimed in any of claims 13 to 24.
44. A non-human animal comprising a genetically modified cell containing a truncated form of a gene sequence substantially as claimed in any of claims 13 to 24.
- 5 45. A toxin formed by processing of a substantially isolated non-toxic precursor polypeptide having an amino acid sequence corresponding to that of an insect specific neurotoxin present in the venom of the Black Widow Spider (*Latrodectus mactans* *Tredecimguttatus*).
- 10 46. A toxin as claimed in claim 45, in which the toxin is formed by truncation toward the carboxy (C) end of the precursor polypeptide.
47. A toxin as claimed in claim 46, in which the toxin amino acid sequence generally corresponds to the amino acid sequence of the precursor polypeptide, truncated by between 150 and 200 amino acids.
- 15 48. A toxin as claimed in any of claims 45 to 47, in which the toxin amino acid sequence is formed by truncation toward the amino (N) end of the precursor polypeptide amino acid sequence.
49. A toxin as claimed in claim 48, in which the fragment cleaved from the amino end is significantly smaller than the fragment cleaved from the carboxy end.
- 20 50. A toxin as claimed in claim 48 or claim 49, in which the fragment cleaved off comprises 7 or 28 amino acids.
51. A toxin as claimed in any of claims 45 to 50, in which the toxin comprises or is an analogue of the insect specific neurotoxin δ -Latroinsectotoxin (δ -LIT), or an active derivative thereof.

52. A toxin as claimed in any of claims 45 to 51, in which the toxin comprises an amino acid sequence as shown in SEQIDN01 and SEQIDN02 or an active derivative thereof.
53. A method of producing an active insect specific neurotoxin polypeptide from an isolated inactive precursor polypeptide having an amino acid sequence corresponding to
5 that of an insect specific neurotoxin present in the venom of the Black Widow Spider (*Latrodectus mactans Tredecimguttatus*), the method comprising truncating the isolated precursor polypeptide.
54. A method as claimed in claim 53, in which the isolated precursor polypeptide is truncated at the Carboxyl end.
- 10 55. A method as claimed in claim 53 or claim 54, in which the truncation is effected using proteolytic cleavage, and preferably by site directed mutagenesis.
56. A method as claimed in any of claims 53 to 55, in which truncation of the N terminus is provided.
57. A method as claimed in claims 53 to 56, in which the active polypeptide is a toxin
15 and is substantially as claimed in any of claims 1 to 12, 45 to 52.
58. An isolated nucleotide base sequence encoding for a toxin precursor polypeptide with an amino acid sequence as shown in SEQIDN04 or a derivative thereof.
59. An isolated base sequence comprising a base sequence as shown in SEQIDN03 or a derivative thereof.
- 20 60. An isolated base sequence as claimed in any of claims 58 or 59, in which the nucleotide base sequence encodes a precursor polypeptide of the neurotoxin δ -Latroinsectotoxin (δ -LIT).

61. An isolated base sequence substantially as provided in the microorganism deposited under Accession No. NCIMB 40633.
62. A recombinant DNA molecule comprising a sequence substantially as claimed in any of claims 58 to 61.
- 5 63. A recombinant molecule as claimed in claim 62, in which the molecule comprises a virus.
64. A recombinant molecule as claimed in claim 63, in which the virus comprises a baculovirus.
- 10 65. A cell, such as a bacterial or viral cell, transformed with a recombinant DNA molecule substantially as claimed in any of claims 62 to 64.
66. An insecticide system comprising means for expressing a base sequence substantially as claimed in any of claims 58 to 61 to produce a precursor polypeptide and to process the precursor polypeptide to produce a toxin in an insect to kill or incapacitate the insect.
- 15 67. An insecticide system as claimed in claim 66, in which the system comprises a viral expression system.
68. An insecticide system as claimed in claim 67, in which the viral expression system comprises baculovirus.
69. A plant comprising a genetically modified cell containing a nucleotide sequence
20 substantially as claimed in any of claims 58 to 61.
70. A non-human animal comprising a genetically modified cell containing a nucleotide sequence substantially as claimed in any of claims 58 to 61.

71. A novel toxin substantially as hereinbefore described with reference to SEQIDN01 and SEQIDN02.
72. A nucleotide sequence substantially as hereinbefore described with reference to SEQIDN01.
- 5 73. An isolated polypeptide substantially as hereinbefore described with reference to SEQIDN03 and SEQIDN04.
74. An isolated nucleotide sequence substantially as hereinbefore described with reference to SEQIDN03.

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C12N15/82 C12N15/63 C12N5/10 C07K14/435
 A01N63/02 A01H1/00 A01K67/027 C07K1/12

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K A01N A01H A01K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>EP,A,0 358 557 (PLANT GENETIC SYSTEMS NV) 14 March 1990</p> <p>see column 1, line 5 - line 31 see column 1, line 50 - column 2, line 60 see column 3, line 6 - line 12 see column 3, line 25 - column 4, line 21; examples 1-4,7,8</p> <p style="text-align: center;">---</p> <p style="text-align: center;">-/--</p>	<p>1-4, 14, 17, 27-30, 33, 37-39, 42, 45, 47, 50, 56, 59, 60</p>

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

Date of the actual completion of the international search

4 July 1995

Date of mailing of the international search report

1 9. 07. 95

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,
 Fax (+ 31-70) 340-3016

Authorized officer

Montero Lopez, B

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>WO,A,91 16433 (PLANT GENETIC SYSTEMS NV) 31 October 1991</p> <p>see page 1, paragraph 1 - page 3, paragraph 3 see page 4, last paragraph - page 6, paragraph 4 see page 8, paragraph 3 - page 12, paragraph 2 see page 13, paragraph 3 - page 16, paragraph 2 see page 17, paragraph 2 - paragraph 3 ---</p>	<p>1-4,14, 17, 27-30, 33, 37-42, 45,47, 50,56, 58-60</p>
X	<p>EUR. J. BIOCHEM. (1993), 213(1), 121-7 CODEN: EJBCAI;ISSN: 0014-2956, 1993 KIYATKIN, N. ET AL 'Cloning and structural analysis of.alpha.- latroinsectotoxin cDNA. Abundance of ankyrin-like repeats' see abstract see page 122, left column, paragraph 4 - page 124, right column, paragraph 1 ---</p>	<p>1-8, 14-19, 47-54</p>
X	<p>FEBS LETTERS, vol. 270, no. 1,2, 17 September 1990 AMSTERDAM NL, pages 127-131, N.I. KIYATKIN ET AL. 'Cloning and structure of cDNA encoding alpha-latrotoxin from black widow spider venom' see page 127, right column, last paragraph - page 128, right column, paragraph 1 -----</p>	<p>1-6, 47-49,53</p>

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0358557	14-03-90	AT-T- 120801	15-04-95
		AU-B- 622702	16-04-92
		AU-A- 4075589	02-04-90
		DE-D- 68922050	11-05-95
		WO-A- 9002801	22-03-90
		EP-A- 0647711	12-04-95
		JP-T- 3503123	18-07-91
WO-A-9116433	31-10-91	AU-B- 639788	05-08-93
		AU-A- 7752691	11-11-91
		EP-A- 0528857	03-03-93